Feeling Reasons
How emotions explain action

Carman, Mary Elizabeth

Awarding institution:
King's College London

FEELING REASONS

HOW EMOTIONS EXPLAIN ACTION

Mary Carman

King's College London

Thesis submitted for PhD in Philosophy

September 2014

ABSTRACT

There is a growing body of literature across disciplines emphasising the way in which emotions are not straightforwardly opposed to reason, as was once typically supposed. In particular, it is often argued that emotions give us evaluative information about the world around us and that they are crucial for the good-functioning of our rational decision-making capacities. Despite this enhanced understanding of the functional role of emotions, however, the extent to which it has implications for our conception of rational agency has yet to be comprehensively addressed. This thesis fills some of that gap.

Part of our conception of ourselves as rational agents is that we guide our actions by reasons, and part of our conception of rational action is that it is action done in light of reasons. In this thesis, I examine what implications an enhanced understanding of emotion has on our conception of rational agency and argue that we can act rationally when acting on the basis of an emotional experience. By examining our concept of 'emotion' and how we explain action via emotion, I argue for four main claims. First, by looking at the role of emotions in explanations of action, I argue that there is conceptual space for emotions to be involved in rational action. Second, I argue that emotions provide access to reasons which could be the reasons in light of which an agent acts and, third, the agent can indeed act in light of them and guide her action when acting on the basis of an emotional experience. Finally, I argue that such action is reasonable, understood as being subjectively rational. My arguments contribute towards a robust conception of rational agency, one which acknowledges our emotional nature and which is able to incorporate emotions into an account of how we do indeed act in many of the cases when we act rationally.

ACKNOWLEDGEMENTS

I would first like to thank Maria Alvarez, my primary supervisor, for her insight, precision, patience and support. My secondary supervisor, Bill Brewer, also provided invaluable feedback and encouragement, and I would like to thank him as well as Clayton Littlejohn and Matteo Mameli with whom I also worked.

I would not have been able to embark on this project without the financial support of the Mind Association and their generous Foreign Doctoral Studentship which I held from 2010-2013. For 2013-2014, I was supported by a Jacobsen Studentship from the Royal Institute of Philosophy and received an award from the King's College London Philosophy Department. I am grateful to all of these organisations for giving me the chance to explore my ideas in comfort.

The King's College London Philosophy Department is a constructive, critical and exciting environment in which to do philosophy. I would like to thank the members of staff and my fellow research students for creating such an environment, and for sharing their expertise and humour in the research seminars, the common room and the Philosophy Bar. I would particularly like to thank the members of the Gender & Philosophy and of the Rationality & Reasoning reading groups, two groups from which I have benefited philosophically and personally. I would like to acknowledge and thank many people individually for their philosophical insight, but shall limit myself to a few notable mentions: Desirée Lim, Janiv Paulsberg, Julie Reid, Maarten Steenhagen and Sasha Vereker.

I have presented papers at various conferences in the United Kingdom, Europe and the United States, where I have received helpful feedback from the participants. I would like to thank the Institute of Philosophy at the University of London, King's College London School of Arts & Humanities and the Philosophy Department for helping me to finance these trips.

On a personal level, there are many friends and extended family who have been central to my life while writing this thesis, and whom I would like to thank for their support. These include Cynthia and Robert Anderson, Francis and Susan Hay, Eva-Maria Burford, Marie-Luise Ermisch, Helen Gibbs, Kathryn Lloyd, Michael Nash and Derek Welton. But I would most like to thank my close family. I have been lucky to grow up with two inspirational grandmothers, Dew Barry and Jean Carman, and a calming brother, Edward Carman. And my parents, Jillian and Nigel Carman: thank you for your continual encouragement, patience, understanding and, of course, your love.

CONTENTS

Abstract **2**

Acknowledgements **3**

Introduction **7**

PART I

1 The richness of the emotion concept **17**
1.1 'Emotion': rich or confused? 18
1.2 What makes an emotion 20
1.3 Conclusion 30

2 Emotions in explanations of action **32**
2.1 Examples of emotions in explanations of action 33
2.2 Conclusion 41

3 Historical overview of emotions and action **42**
3.1 Historical overview 43
3.2 Twentieth-century developments 47
3.2.1 Ryle on causes and motives 49
3.2.2 Kenny on motive patterns 53
3.2.3 Different ways in which we can act from an emotion 57
3.3 Conclusion 63

PART II

4 Accounts of emotion **67**
4.1 Views which foreground non-cognitive elements 69
4.1.1 Bodily changes and their feelings 71
4.1.2 Accommodating the quality of feeling 73
4.1.3 Prinz and embodied appraisals 75

4.2 Views which foreground cognitions 80
4.2.1 Belief-desire accounts 81
4.2.2 Judgement accounts 83
4.3 Weak cognitive accounts 88
4.3.1 Different kinds of cognitivism 88
4.3.2 Perceptual models 89
4.4 Conclusion 95

5 Rationalising explanations 96
5.1 Rational action 98
5.2 Beliefs, desires and motivating reasons 101
5.2.1 Explaining an action in terms of beliefs and desires 102
5.2.2 A problem for motivating reasons 107
5.3 The limit of belief-desire explanations 111
5.3.1 First problem: mischaracterising & trivialising the action . . 113
5.3.2 Second problem: trivial explanations 117
5.4 Motivating reasons and explanatory reasons 119
5.5 Conclusion 124

PART III

6 The intentional nature of emotions 128
6.1 The formal object of an emotion 129
6.2 The requirements of the formal object 133
6.2.1 What the role of the formal object commits us to 134
6.2.2 Counterexamples to the individuating role 135
6.2.3 Clarification on the intentional nature of emotion 141
6.3 Using the formal object 142
6.3.1 The formal object in assessments of intelligibility 143
6.3.2 The formal object in assessments of accuracy 145
6.3.3 The formal object in assessments of appropriateness 147
6.3.4 A final example 147
6.4 The emotional reasons 149
6.5 Conclusion 150

7 Acting reasonably in light of a reason 152
7.1 Acting in light of a reason when acting on the basis of an emotion . 154
7.1.1 Döring's argument 154

7.1.2 Meeting the guidance condition and acting reasonably . . . 162
7.1.3 Acting in light of a reason 166
7.2 Acting reasonably on the basis of an emotional experience 169
7.2.1 Emotional experiences as evidence and reasons 171
7.2.2 Emotional experiences as providing prima facie entitlement to act 180
7.3 Conclusion 186

Conclusion **191**

References **195**

INTRODUCTION

> *That there is no more miserable Slave than he who suffers to be guided by his passions.*

This is how Jean Senault titles the Third Discourse in his *The Use of the Passions* (1649). What follows is an extended metaphor comparing the passions to the worst kind of slave owner and ourselves to the suffering slaves. Being enslaved and deprived of freedom, Senault tells us, is 'not comparable to that which the tyranny of the Passions causeth in us' (95). Nevertheless, he goes on to argue, the Passions contain the seeds for both virtue and vice and he concludes that 'there is no Passion in our Soul, which may not profitably be husbanded by Reason and by Grace' (510).

Senault, and others who oppose emotion to reason, are too heavy-handed in their treatment of emotions. Our understanding of emotion in both psychology and philosophy has increased substantially in the past century and it now seems by no means clear, or correct, that the most 'miserable Slave' is 'he who suffers to be guided by his passions'. It is by no means clear, or correct, that being guided by our emotions necessarily deprives us of acting in line with our reasons if our emotions are not 'husbanded by Reason'. In this thesis, by focusing on explanations of action, I argue that we can act rationally when acting on the basis of an emotion, in the emotion's own right. We are not miserable slaves to our emotions.

In the paper 'Quick and Smart? Modularity and the Pro-Emotion Consensus' (2008), Karen Jones gives the label 'the pro-emotion consensus' to the school of views in both psychology and philosophy which rejects the opposition of emotion to reason. Those who form part of the pro-emotion consensus reject the claim that emotions are 'at best arational, if not outright irrational' (Jones 2008, 3), needing to be 'husbanded by Reason' as Senault puts it. They instead see emotions 'as evolved capacities that are integral to our practical rationality' (3), and which can even 'key us to the presence of real and important reason-giving considerations' (Jones 2003, 181), considerations constituted by how the agent's interests and concerns are affected in a given situation. As Jones writes:

> Pro-emotion theorists also think that their position has revisionary implications, perhaps even radical ones. There is a tendency within philosophical and commonsense thinking to disparage the emotions and to suppose that wise deliberation and objective inquiry are dispassionate. But if emotions contribute positively to ecologically situated rationality, then commonsense and philosophical norms of rationality that embed false assumptions about their disruptiveness may need to be revised. (Jones 2008, 4)

While Jones classes herself as a pro-emotion theorist, she does warn that the revisionary implications of the pro-emotion school need careful defence.[1] I too place myself amongst the pro-emotion consensus and my aim in this thesis is to provide some of that defence. My focus is on the questions: if emotions are crucial to our functional rationality, and if emotions are intentional states which provide us with information, how should we understand their contribution to our capacity as rational agents who act for reasons? And what kind of explanation can we give of action performed because of an emotion?

Two key developments in particular form the backbone of the pro-emotion consensus school of views and cast doubt on the traditional emotion-reason opposition. These are: first, the idea that emotions are not simply affective states which only play a motivating role in action; they also give us information about the world around us. Second, emotions are crucial for the good-functioning of our rational decision-making capacities.

The first development is the idea that emotions provide us with information, and has been developed in both psychology and philosophy as the claim that emotions are intentional states and are about things other than our bodies.[2] In particular, emotions are said to represent and be elicited by evaluative properties relating to how we stand in our environment. This gives rise to the challenge of explaining how emotions represent and are elicited by such properties without themselves being composed of beliefs, judgements, or other cognitively complex states. Assuming that such a challenge can be met, it seems that emotions are *prima facie* states which give us evaluative information about how things are, and we can plausibly ask whether someone can act rationally from an emotion and, by extension, what role an emotion can play in an explanation of rational action.

[1] Jones' own focus is on whether emotions, as adaptive mechanisms of the mind, 'are capable of coming to be directed towards new objects in virtue of a cognitively modifiable range of triggering properties' (Jones 2008, 4).

[2] For instance, appraisal theories in psychology develop the idea that emotions involve appraisals or dimensions of appraisals relating to the subject and her environment. See Agnes Moors (2014) for a concise description of different appraisal theories.

The second key development is the idea that emotions are crucial for the good-functioning of our decision-making capacities. Emotions facilitate planning by helping to organise goals, and they assist deliberation by giving certain considerations salience, thereby ensuring that we can function as rational agents, capable of deliberation, decision-making, and acting. The neuroscientist Antonio Damasio (1994), for example, famously proposes a somatic-marker hypothesis for the role of emotion in decision-making. The somatic marker hypothesis is the hypothesis that emotions, as bodily states, affectively mark aspects of a situation or potential outcomes as positive or negative. This allows the agent to efficiently assess the pros and cons of the potential outcomes, both enabling decision-making and, via a feedback loop, ensuring that a person can decide quickly without always having to go through slower, deliberate processing.[3]

We also find dual-process theories of mind, made especially popular with Daniel Kahnemann's best-selling book *Thinking, Fast and Slow* (2011) and recently defended by Jonathan Evans and Keith Stanovich (2013). These theories distinguish between rapid autonomous processes (Type 1), which would include emotions, and higher order reasoning processes (Type 2), and often make the further claim that we use Type 1 processing more than we think. As Kahnemann writes, 'the intuitive [Type] 1 is more influential than your experience tells you, and it is the secret author of many of the choices and judgements you make' (Kahnemann 2011, 13).[4] This has been applied to philosophical topics, for example by Jonathan Haidt in a paper titled 'The Emotional Dog and its Rational Tail' (2001), who argues that moral judgements are generally the result of quick, automatic intuitions and not the result of moral reasoning. Moral reasoning, he argues, is post hoc. Drawing on Haidt's work amongst others, the psychologist Joshua Greene argues that our characteristically deontological moral judgements are driven by automatic emotional responses, whereas our characteristically utilitarian moral judgements, which require more reflection, are driven by more controlled cognitive processes (Greene 2007).

[3]How far the somatic marker hypothesis can be used to argue for a central role of emotion in decision-making has been challenged. For example, Stefan Linquist and Jordan Bartol (2013) argue that the evidence only supports emotions having a peripheral role, either 'in notifying the subject that a decision-point has been reached', or 'at the terminal stage, in motivating the subject to execute a decision' (2013, 2). Nevertheless, the evidence is still sufficient to suggest that emotions can notify the subject of information and motivate her to act in a way independent of conscious deliberation. As Linquist and Bartol are happy to concede, more cognitively complex models of emotion could play a more central role in decision-making.

[4]See Magda Osman (2013) for a critique of dual-process theories. Osman argues that the evidence is ambiguous and could support a quantitative rather than qualitative distinction between Type 1 and Type 2, so that 'the same reasoning process generates accurate responses to easier tasks and generates inaccurate responses to harder tasks' (Osman 2013, 251).

Whether or not such interpretations and applications to moral reasoning are successful, there is still sufficient evidence that emotions assist with reasoning by focussing attention and speeding up decision-making, as well as providing an alternative route for coming to a decision, a route that bypasses more deliberative reasoning.

In their contribution to the functioning of human reasoning, emotion and deliberation need not be in opposition, where the latter is usually what is understood as reason. Nevertheless, how the underlying processes work need not have a direct bearing on our conception of rational action. To see this, compare goal-directed behaviour with rational action. Goal-directed behaviour is, quite simply, behaviour that is aimed at achieving some goal. Examples would include my going to the shops to buy milk for my cereal, a dog's scratching of the cupboard door to get food, and, arguably, even a sunflower's turning its head towards the sun. Emotions undeniably play a role in goal-directed behaviour: they function as a mechanism through which we decide and act timeously in order to achieve our goals.

Rational action is something quite different. I want milk for my cereal and so I go to the shop in order to buy milk. This is a clear example of rational action. It is controversial whether animal goal-directed behaviour is also rational, but it is uncontroversial that the sunflower does not act, let alone act rationally. Rational action is thus not straightforwardly the same as goal-directed behaviour. Emotions may assist with goal-directed behaviour by contributing to the functioning of rational decision-making, but it does not follow that they thereby have a crucial role to play in our conception of rational action. That emotions can contribute to the good-functioning of rational decision-making is therefore not enough to establish what role they have, if any, in our conception of rational action.

The significance of my thesis is that it bridges part of the gap between our understanding of the functional role of emotions and our conception of rational action. I focus on explanations of action because explaining an action as rational involves showing that it is rational, so focusing on explanations of action allows us to examine our conception of what rational action is. By examining our concept of emotion and how we explain action via emotion, I argue that, because emotions provide access to reasons in light of which we can act, we can act in light of a reason, hence act rationally, when acting on the basis of an emotion. If I am right, then we end up with a robust conception of rational action which is able to incorporate emotions into an account of how we often do act rationally. Even if we do sometimes guide our actions by emotions, we need not be miserable nor slaves, controlled by evil masters.

STRATEGY AND STRUCTURE

In order to argue for the view that we can act in light of a reason when acting on the basis of an emotional experience, my thesis is divided into three parts. In Part I, chapters 1-3, I argue that there is conceptual space for emotions to have a reason-providing role in action and in explanations of action. In Part II, chapters 4-5, I give necessary background for formulating the kind of cases in which I am interested as well articulating the view I shall develop. In Part III, chapter 6-7, I develop the details of the view.

The overarching argument of Part I is to take two uncontroversial claims and, by showing how they interact, motivate the case for the much more controversial thesis that acting on the basis of an emotional experience can be acting in light of a reason.

The first claim that I develop in Chapter 1 is that our concept of emotion is rich. We refer to a diversity of different mental states or attitudes as emotions or by using emotion words, such as referring to short-term agitations, occurrent emotion states, long-term attitudes, even moods. It is also rich because our concept of emotion is of something that is both intentional and has a specific phenomenology. In Chapter 1, I emphasise the intentional nature of emotions and introduce the concept of the formal object of an emotion, a concept which will play a crucial part in developing the view that emotions provide access to reasons. The formal object is the property in virtue of which a particular emotion is elicited and is the emotion that it is. For example, if I am afraid of the dog, the dog has the property of being dangerous or frightening, a property in virtue of which my emotion is elicited and is fear. The formal object is often said to be what the emotion is about; in this example, my fear would be about the dangerous or frightening dog.

The second claim that I develop in Chapter 2 is that we can give a range of explanations of action involving emotions, and the role of the emotion in the explanation need not be the same. I describe five categories of different roles that the emotion could play in an explanation and give examples to illustrate each category. The examples start with those where the role of the emotion is tangential to the action *qua* rational. Gradually, the emotion's involvement with the action *qua* rational increases, ending with examples where the emotion is plausibly the agent's reason and examples where the emotion plausibly provides access to the agent's reason for acting through its intentional nature.

An emotion could play a role in an explanation of action in virtue of being an agitation, an occurrent state, an attitude or a mood; it could play a role in virtue

of its phenomenology or in virtue of its intentional nature; it could even play more than one role. We therefore see that, even if an emotion plays one kind of role in one kind of explanation, it need not play the same role in all explanations. Despite this, emotions have traditionally been relegated to only one or perhaps two roles. In Chapter 3, I show this by presenting representative themes of historical approaches to emotions and action, and by focusing on two twentieth-century accounts of emotions in action and explanations of action in more detail, those of Gilbert Ryle and Anthony Kenny. Both Ryle and Kenny argue that emotions have one or two roles in action and action explanations: either as causes or as motives. I argue that their accounts are not satisfactory because they cannot accommodate the differences between the categories of examples that I introduced in Chapter 2. They either do not adequately incorporate the intentional nature of emotion into their account, as in the case of Ryle, or they do not consider the full impact of the intentional nature, as in the case of Kenny. I conclude that, if we properly embrace the intentional nature of emotions, one of the roles available to emotion in an explanation of action could be in virtue of its giving the agent's reason for acting as she did, and the conceptual space for acting in light of a reason when acting on the basis of an emotion is opened up. This brings the first part of my thesis to a close.

In Part II, I cover important background for articulating a basis from which I can develop my positive view that we can act in light of a reason when acting on the basis of an emotional experience. The view turns heavily on our conception of emotions as intentional states. So, in Chapter 4 and again in Chapter 6, I examine emotions and their intentional nature in more detail. In Chapter 4, I focus my attention on occurrent emotion states and address what an adequate account of emotion must look like if it is to capture accurately their intentional nature. I examine two main groups of accounts of emotions, non-cognitive and cognitive accounts. I argue that non-cognitive accounts are unable to accommodate the intentional nature in a way which does not undermine their core claims, and that cognitive accounts which analyse emotions as involving beliefs of judgements fare no better as they either over-intellectualise emotions or are trivial. I therefore propose that we should adopt a weak cognitive account, such as the currently popular perceptual models of emotion. On such accounts, emotions are understood as intentional states that are about the formal objects, yet are not cognitively complex.

A final bit of background work is required, which is to lay out what rational action is and what an explanation of rational action *qua* rational action requires. I argue that acting rationally is acting in light of a motivating reason, where a motivating reason is a fact. I then argue that we can explain an agent's action by identifying

the motivating reason and argue for a distinction between explaining an action as rational, or done for a reason, and explaining it as reasonable. Both acting rationally and acting reasonably are important for our conception of ourselves as rational agents. I end by describing more formally what the differences between the categories of explanation in Chapter 2 are, and stating more precisely how, because of their intentional nature, emotions could plausibly provide access to the agent's motivating reason and how an explanation citing an emotion could be an explanation of the action as reasonable.

In Part III, I develop the details of the view that we can act in light of a reason when acting on the basis of an emotional experience. In Chapter 6, I return to emotion's intentional nature in order to establish what, exactly, the agent's reason could be by examining the concept of the formal object, which I first introduced in Chapter 1. I argue that our concept of the formal object is a response-dependent concept where one of the elements is a relational property and another element is the subject as actively caring. For example, the formal object of fear is the property of being frightening and being frightening is to merit fear. Part of being frightening is being dangerous and it is in virtue of the subject having the active cares that she does that the property of being dangerous elicits fear. While an emotional experience as a whole will represent the formal object, I argue, an emotion itself need not be explicitly about the formal object. So, my fear of the dog represents the formal object of the dog's being frightening, but my fear may only be about the dog as dangerous. From this, I propose that there are three options for the agent's reason: the fact of how things seem to the agent, e.g. the frightening dog (to me); the fact of how things are, as in what response is merited, e.g. the frightening dog; and the fact of how things are, as in what relation holds, e.g. the dangerous dog.

In Chapter 7, I expand on my positive view by responding to two brands of objection. The first brand is that we cannot act in light of a reason when acting solely on the basis of an emotion because we need to act on the basis of a belief or judgement. By drawing on the concept of the formal object as response-dependent, I argue that, through being emotionally committed to how things appear, we are in the right kind of state to act in light of the reasons which emotions provide. The second brand is that we cannot act reasonably when acting solely on the basis of an emotional experience. By referring to the three options for reasons that I identified in Chapter 6, I argue that we can act reasonably because emotions do provide evidence of how things are, albeit of different strength depending on the reason for which it is evidence, and that we have a *prima facie* entitlement to act on the basis of an emotional experience. I therefore conclude that we can act in light

of a reason when acting on the basis of an emotional experience and that we can explain an action as reasonable by citing an emotion.

IMPORTANT CONTRIBUTIONS OF THIS THESIS

The significance of this thesis can be seen in three of the important contributions it makes which arise from my specific focus on explanations of action.

The first important contribution is that I provide a comprehensive analysis of a range of roles that emotions can play in an explanation, from being independent of explaining the action *qua* rational through to being fundamental to the explanation of the action *qua* rational. The role of emotions in explanations in the existing literature has been addressed piecemeal, by looking at particular kinds of explanations such as explanations of expressive action, or proposals that argue that emotions play only one or perhaps two roles. In contrast, my argument turns on identifying a range of roles and comparing them to each other in order to draw out both what our concept of emotion and what our concept of rational action require. This contribution is important because it creates order amongst different ways in which emotion can be involved with rational action, but also because it can potentially be drawn on to resolve debates about the epistemic status of reasons for belief, a contentious topic in the philosophy of emotion. This is because one lesson from my thesis is that emotions can be involved with reasons for action in a number of ways, and the same may be true for reasons for belief.

The second important contribution is that I examine emotions and action using the framework of reasons as facts, where acting in light of a reason is acting in light of a fact. Examinations of emotions and action have tended to use the framework of reasons as mental states. Using an alternative framework not only fills a lacuna in the topics surrounding emotions and action, it also brings out distinctions which might otherwise be overlooked, such as a distinction between acting on the basis of an emotional experience where the emotion, as an experience, is the reason in light of which the agent acts, and acting in light of the reason an emotion provides access to. This enriches our understanding of the emotional landscape.

The third important contribution is that I argue that we can act in light of a reason when acting on the basis of an emotion without requiring a mediating belief or judgement. Much work has been done on the epistemic status of emotions as being or providing reasons for belief or knowledge. Rather than basing the reason-providing potential of emotions on their epistemic role pertaining to beliefs, I address straight-on emotion's role in action. This is an important contribution because

it addresses action directly and, while arguments against the reason-providing role of emotions for belief may also count against my view, they do not do so automatically. Examining how the arguments apply could potentially feed into other areas of inquiry, such as the relation of reasons for belief to reasons for action.

There are other contributions which my thesis makes, but they are related to specific arguments or claims I develop, and how I see further research in light of these arguments and claims. I shall therefore rather say something about these contributions in my conclusion once those arguments and claims are familiar.

So, I now turn to the first part of my thesis and argue that there is conceptual space for emotions to have a reason-providing role in explanations of action.

PART I

CHAPTER 1

THE RICHNESS OF THE EMOTION CONCEPT

Our concept of 'emotion' is rich: we use emotion words to refer to a diversity of things. We might think an occurrent instance of fear is an emotion, but also a long-term attitude of love, or even a sad mood. We might think of an occurrent state as fear, but we also sometimes think that someone can be fearful or act out of fear even when she is not in a state of fear at the time. In terms of this diversity, the concept of 'emotion' is perhaps similar to our concept of 'object', which we also use to refer to a diversity of things, including physical items as well as abstract 'objects' of thoughts. But then, just like we might argue over whether something really is an object and worry about the concept of 'object', we might argue over whether something really is an emotion and worry about the concept of 'emotion'. We might worry, that is, that our concept is confused rather than rich.

My aim in this chapter is to show that the concept is rich. In section 1.1, I begin by introducing examples of emotions, illustrating ways in which they vary and pointing out how there is a lack of consistency even in what theorists use the very word 'emotion' to refer to. I therefore need to show that the concept of emotion is not confused as well as define how I shall be using terminology. This is what I do in section 1.2. I distinguish between agitations, occurrent emotion states, emotional attitudes and moods. Despite their differences, I show that they are relevantly similar because they share two main features to greater and lesser degrees. These features are: (i) that emotions have an intentional nature that is evaluative; and (ii) that emotions have an affective nature including an intentional quality, a distinctive phenomenology based on their evaluative intentional nature. These two features are what the diversity of things we refer to as emotions have in common and they are what make the concept rich rather than confused.

1.1 'EMOTION': RICH OR CONFUSED?

What are emotions? Despite our personal familiarity with emotions, defining what emotions are is a difficult and controversial task. Indeed, even amongst the specialists investigating emotions in a field like psychology, there is a notorious lack of consensus.[1] It might be best to start with some examples.

Love, happiness, fear, anger, sadness, disgust, shame, guilt, jealousy, joy, surprise. These are just some of the things we think of as emotions. Even within this list, there is variety. The kind of object of the emotion varies, for example. You may love your partner or you may love an ideal; you may fear the intruder or you may fear the uncertainty of the future. In these cases, the first of the pair is love or fear of a concrete object while the latter is love or fear of something abstract.

Emotions also vary in complexity, in what they require in terms of human development or in terms of their internal structure. For instance, fear, anger, sadness and disgust are examples of the so-called 'basic emotions'. These are emotions which are claimed to characteristically occur universally across human cultures, appear early on in human development, and result from an innate capacity. They are also often claimed to be modular, processing information regarding a single domain.[2] Other emotions, like jealousy and pride, appear later on in human development because, in order to be jealous of another or proud of oneself, the person needs to have developed a sense of self as well as a 'theory of mind', which is required in order to attribute mental states to oneself and to others.[3] Some theorists even propose that non-basic emotions arise out of or are blends of basic emotions with additional

[1] In *Emotion Review*'s (2010) special section on 'Defining Emotion', the psychologist Carroll Izard presents a survey he did of 35 emotion scientists asking them questions about definitions. His conclusion: 'there is still no consensus on a definition of "emotion", and theorists and researchers use "emotion" in ways that reflect different meanings and functions' (Izard 2010, 363).

[2] The psychologist Paul Ekman's original list of basic emotions, devised on supposedly universal facial expressions and used extensively in psychological research on the expression of emotions, includes fear, anger, sadness and disgust as well as happiness and surprise. He later considers adding contempt, shame, guilt, embarrassment and awe to the list (Ekman 1992). For a useful table summarising a sample of theorists defending basic emotions, what emotions they include and on what basis, see Andrew Ortony and Terence Turner's 'What's Basic About Basic Emotions?' (1990, 316). Ortony and Turner take the lack of agreement as a starting point to argue that there is no coherent and non-trivial notion of a basic emotion.

[3] Michael Lewis, for example, has worked on the emergence of a sense of self in infants and shown that, before children have a sense of self against which they can evaluate how they meet some standard, they feel happiness in response to successfully achieving a goal. Once they develop a sense of self, they are more likely to feel pride (Lewis 2010, 307).

beliefs or evaluations.[4]

What we refer to when we speak about individual emotions also varies. For example, walking happily down a dark and empty street, I round a corner and come face-to-face with a figure holding a knife. My fear is immediate, occurrent and embodied. In contrast, my fear may be more mood-like as when, after watching a zombie film and walking down the same dark street, I am jittery and see every shadow as a potential threat, my heart racing in fearful anticipation.[5] My fear could even be a standing attitude. Someone with a fear of flying, for instance, is afraid of flying even if she does not feel the occurrent fear which I feel as I walk down the street. Her fear may become apparent in the kinds of things she says and does, such as expressing dismay at the prospect of flying and choosing alternative means of travel. Despite these differences, we do describe the occurrent state, the mood and the standing attitude with reference to the same emotion, fear.

In all of these cases, the emotion is conscious, but we can also make sense of unconscious emotions. You may consciously feel anger on being cut-off in traffic but you may ignore it, bottling it up. Someone undergoing psychotherapy may discover that she has a deep-seated fear generated from an unpleasant and forgotten event in her childhood which influences her actions and attitudes.

These variations in complexity and variations in individual cases of emotion might influence what we call an emotion. For example, the psychologist Paul Ekman defines emotions as the basic emotions, mentioned above, and argues that emotions only last a couple of seconds. He distinguishes emotions from emotional attitudes and emotional plots. An attitude, he writes, is 'more sustained, and typically involves more than one emotion' (Ekman 1992, 194). If you have an attitude of love, say, you may experience the basic emotions of happiness and fear. Emotional plots are even more complex and 'involve settings and stories in which emotions occur' (194). Things like grief or jealousy are examples of plots. Grief, for instance, describes the context in which one feels the emotion of sadness, namely a context in which someone has died and not merely one where there has been a loss. Ekman also distinguishes attitudes and plots from emotional traits, such as being hostile or melancholic, and from emotional disorders, like depression or mania. Emotional traits and disorders resemble the kinds of things we often refer to as moods.

[4]Robert Plutchik (2001) proposes the hypothesis of non-basic emotions as blends of basic emotions, and Keith Oatley and Philip Johnson-Laird (1987, 44-47) discuss remorse as arising out of sadness, together with an evaluation of having lost some sense of self.

[5]We also refer to character traits with emotion words. If I am always taking fright, for example, we may refer to my character as being a fearful (or cowardly) one. I shall not be concerned with character traits in this thesis.

The philosopher Peter Goldie has a completely different set of distinctions and labelling. In *The Emotions: A Philosophical Exploration* (2000), he distinguishes between an emotion and an emotional episode. An emotional episode involves different elements, such as perceptions, thoughts, feelings and bodily changes. For Goldie, an emotional episode is a particular experience of an emotion, and an emotion is something that is more enduring than a particular episode. It is typically complex, episodic, dynamic and structured. It is complex because it involves many things, such as the perceptions, thoughts, feelings and bodily changes that make up a particular episode. It is complex because it also includes past experiences and dispositions to have emotional experiences, or to have other thoughts and feelings, or to behave in certain ways. An emotion is episodic and dynamic because the different elements come and go over time. It is structured because what holds all of these elements together as one emotion is that they are embedded within a narrative which provides structure to the experiences.

There is therefore a great diversity amongst what we label as emotions. Just comparing Ekman and Goldie, we also see a lack of consistency in how even the word 'emotion' is used. For Ekman, 'emotion' refers to a short episode while for Goldie it refers to an enduring collection of episodes, more like Ekman's attitudes and plots. This diversity and lack of consensus could lead us to reject the idea of emotion as a natural kind (Griffiths 1997) or the concept of 'emotion' as a unitary concept (Izard 2010). For the purposes of scientific study, these differences can relevantly change the target of research. However, I am not investigating what emotions 'really' are but what role they play in action and explanations of action. Even accepting that there are fundamental differences amongst the things we class as emotions, my interest is in why we still refer to them as emotions at all, and whether these different things we think of as emotions play similar or different roles in action and action explanation.[6] My task in this chapter, then, is to identify why we do refer to them as emotions of one kind or another and to bring some clarity to what the concept of an emotion involves.

1.2 What makes an emotion

In the rest of this chapter, I examine what is shared by the diversity of things which we refer to as emotions, as well as define how I shall be using terminology. I begin

[6] Julien Deonna and Fabrice Teroni also explore the diversity and unity of emotions and, as they suggest, 'understanding the thoughts, actions and feelings of others' may even 'require . . . the introduction of categories that do not correspond to natural kinds' (Deonna and Teroni 2012, 23).

by looking at some examples which compare emotions with non-emotions, and identify what makes the emotion different from the non-emotion. I shall conclude that the key differences are the kinds of feeling involved in an emotion and the way an emotion is about something. I then elaborate on these identifying features by examining variations within emotions and variations across emotions, as well as using the features to identify differences between emotions and moods.

To start, take three sets of examples comparing emotions with non-emotions:

EMOTION VS. NON-EMOTION:

1. A. Jesse apologises because he feels guilty.

   B. Jesse apologises because his mother told him he must.

2. A. Jo has a flushed burning face and feels angry.

   B. Jo has a flushed burning face and feels sunburnt.

3. A. Nat is compulsively tapping her fingers because she is excited.

   B. Nat is compulsively tapping her fingers because the scientist is inducing a reaction in her brain.

Probably the most common understanding of emotions is as feelings. You feel angry, afraid, happy, sad, affectionate. Without saying anything further about what these feelings are, we can already see that the difference between 1A and 1B is that Jesse apologises because he feels a certain way in 1A but we have no idea whether he feels anything at all in 1B.

But we do need to say more about what the relevant feelings are. Some emotions involve strong bodily sensations, like feeling your heart race out of love or your palms sweat out of fear. Others do not, like having a long-term fear of flying. Even for those emotions which do have strong bodily sensations, the bodily sensation does not exhaust the way the emotion feels. Jo, for example, has the same bodily sensation of a flushed burning face in both 2A and 2B, but in 2A that bodily sensation does not exhaust her feeling of anger.[7] Emotions have a phenomenology that goes beyond bodily sensations.

One of the particular features of the way emotions feel is that they characteristically have a distinct valence to the way they feel and we often divide emotions

[7] I shall provide support for this claim in Chapter 4 when I critically examine accounts of emotion which hold that emotions are just feelings of bodily changes.

according to whether they are 'positive' or 'negative'. Examples of positive emotions would be love, pride, admiration. Examples of negative emotions would be fear, anger, regret. As Julien Deonna and Fabrice Teroni describe, talk of valence tends to be cashed out in one of two main ways (Deonna and Teroni 2012, 12-14). The first is in terms of 'hedonic' quality and what the emotional experience is like. The second is in terms of conative states, like motivational tendencies or desires. They conclude later in their book that there is a variety of ways in which an attitude or state could be classed as positive or negative, something which Robert Solomon (2003) also argues, and we need to beware of forcing emotion-types and emotion-tokens into one valence and one valence only. Self-righteous anger, for example, can be hedonically pleasant and positive, but can also be unpleasant and negative. Nevertheless, emotional experiences tend to have certain positive and negative feelings to them, regardless of how those feelings arise or whether we can classify emotions on the basis of distinct valences.

What does seem to be the case, however, is that whether something is pleasant or not, or what desires and motivations one has, will be related to what an emotion is about. When we say that Jo feels angry, we are alluding to the way that her feeling is *about* something beyond the way her body has changed. Anger is not simply a burning face and the feeling of anger is not just of the bodily sensation. Returning to the valence, we see that her anger is negative partly because it is about something that has caused her anger, something that has caused her offence in some way. In the case of self-righteous anger, in contrast, what seems pleasant, enjoyable or positive about being in such a state is that one sees oneself as superior. What is unpleasant, not enjoyable or negative is that something is nevertheless offensive.

The idea that emotions feel a certain way and are about something also helps us to identify a difference between 3A and 3B. We do not understand that Nat taps her fingers in 3A simply as the result of some neurochemical change, although it may very well be the case as well. What we understand is that she is tapping her fingers out of excitement, that she *feels* excited and that there is something she is excited about. In 3B, there is no such implicit reference.

Much more needs to be said about what emotions are about and how this relates to their phenomenology, which I shall do by introducing four more sets of examples which illustrate variations within emotions, across emotions, and between emotions and moods.

VARIATIONS WITHIN EMOTION:

4. A. Alex ruffles his daughter's hair because he feels a surge of love.

   B. Alex always makes sure he is home early from work in order to spend time with his daughter because he loves her.

5. A. Kim fears the figure with a knife.

   B. Kim fears his next Actuary exam.

VARIATION ACROSS EMOTIONS:

6. A. Sasha hurried home because she feared the rottweiler.

   B. Sasha hurried home because she loved the rottweiler.

EMOTION AND MOOD:

7. A. Sam was sad about the death of her cat.

   B. Sam was generally feeling sad and low.

Peter Hacker's framework in 'The Conceptual Framework for the Investigation of Emotions' (2004) is useful here for fleshing out the differences illustrated by these examples.[8] Hacker distinguishes between agitations, emotions and moods and argues that they lie along a continuum. I shall adopt his framework and, by discussing how these differ from each other, also expand on what emotions are about and the kind of feelings that they have.

Agitations, for Hacker, are 'short term affective disturbances, typically caused by something unexpected' that we perceive, learn or realise (Hacker 2004, 201).[9] Agitations are fundamentally reactive and are things like feeling excited, thrilled, shocked, revolted, disgusted. They have bodily manifestations and, in this sense, are bodily events with associated sensations. Agitations lie on one extreme of the continuum. Next are emotions.

Within emotions, Hacker distinguishes between occurrent emotion states, what he calls 'perturbations', and emotional attitudes. Occurrent emotions are closer to agitations than are attitudes but, because they fall along a continuum, there is not

[8]Hacker does not aim to provide necessary and sufficient conditions for what an emotion is; rather, he aims to tidy up our concept and provide a framework that picks out some of the important conceptual similarities and differences. This is exactly what I aim to do in this chapter and so it is a useful framework to draw on.

[9]In terms of their length, Hacker's 'agitations' correspond roughly to Ekman's 'emotions'.

a clear distinction between occurrent emotions and agitations. Like agitations, occurrent emotions have a strong affective element often made up of the sensation of some bodily manifestation. Like agitations, occurrent emotions can result in expressive behaviour, such as blushing. They can also affect the manner in which we act, such as a tone of voice. Unlike agitations, however, they can result in motivated action such as retreating out of fear rather than reactively jumping backwards.

To illustrate the difference between agitations and occurrent emotions, let us expand on example 5A where Kim feels fear when coming face-to-face with a figure brandishing a knife. Suppose he reacts in fear and jumps, but very quickly realises that the figure is actually his friend with an umbrella. His agitation dissipates and has no lasting effect on his action choices. If he reacts in fear and jumps, but continues to feel perturbations and is afraid, then he is in an occurrent state of fear which could affect the kinds of things he chooses to do, such as to continue retreating backwards. We could also see Jo's flushed burning face in 2A as an example of an agitation, but 3A, 4A, 6A and 6B as examples of occurrent emotions.

Next along the continuum are emotional attitudes. Examples of attitudes would be feeling long-term love for a person, ideological hatred for the political opposition, or fear of flying. Agitations and occurrent states are similar in terms of being reactions and involving bodily changes; they both characteristically involve bodily sensations. Attitudes, in contrast, do not necessarily have bodily manifestations and associated sensations; neither are they necessarily dispositions to have such manifestations or to experience the corresponding occurrent emotion state.

To see this, consider love for one's parents. I love my parents right now as I write and have done so my whole life. However, I am not in a permanent physical state that corresponds to the occurrent state of love or affection. Neither am I disposed to experience the occurrent emotion of love with its accompanying bodily manifestations in response to them. When I am with them, I do not feel my heart racing or feel agitated exuberance, feelings of the occurrent state of love. If anything, I am disposed to feel a range of other emotions: pride when they do something admirable, anger when they do something silly, distress when they are in danger. Being disposed to feel occurrent states of love does not exhaust why I can truthfully say that I love my parents, or why I love them even when I am angry, or even why my love disposes me to anger.

Instead of being dispositions to have occurrent emotions of love, the attitude of love is, as Hacker insightfully writes, 'a *lasting concern* for the object of love, a *standing motive* for action beneficial to the beloved, a *desire for shared experiences*, and a *persistent colouring* of thought, imagination and fantasy' (Hacker 2004, 203,

my emphases). I have a lasting concern for my parents: amongst other things, I worry about them when they are sick or when I have not heard from them; I get angry when I think that they have been offended in some way. I have a standing motive for action beneficial to them: I carefully pick out birthday presents and pick up the groceries when I realise that they will not have a chance to do so themselves. I desire to be with them: I miss them when I am away and try to go home to see them as often as possible; I get upset when I feel that they are not doing the same for me. My thoughts are persistently coloured with their value to me: even when angry, I see them as worth being angry with, I make decisions about where to live with them featuring as considerations to weigh up, and I get distressed if I imagine a life without them. In this way, Hacker's attitudes are like Ekman's attitudes and plots, and also like Goldie's emotion. They should not be understood as a discrete attitude so much as a collection of behavioural tendencies, attitudes, occurrent emotions, etc., held together by an overarching theme of what would make it love or fear or anger.

Many emotions are standing attitudes rather than occurrent states. The person who is afraid of flying and avoids flying altogether, for example, may be disposed to experience occurrent fear but, because she never flies, that fear is never made manifest. She may even be disposed to feel excitement whenever her colleagues organise alternative means of travel for business trips. Yet, she does genuinely fear flying. Amongst other things, this is seen in her decision to give up flying: she has a lasting negative and fearful attitude towards the prospect of flying, a standing motivation to travel by any other means, and a persistent colouring of her thoughts of flying. We can now distinguish between 4A and 4B. Example 4A is of Alex's occurrent state of love and 4B is of Alex's long-term attitude of love. His love affects his choices and he strives to spend time with the person he values.

If attitudes and occurrent emotion states are so different in terms of bodily manifestations, what makes them both emotions? There is a point of similarity which is that both occurrent states and attitudes are about something. They are intentional.

'Intentional' is a term of art in philosophy. I take 'intentionality' to relate to how a mental state is about, directed or targeted at objects, and how it represents those objects as being certain ways. Intentional states or attitudes have an intentional object, which is what it is about or targeted or directed at, and an intentional content, which is how the object is represented as being. A typical intentional state is a belief. If I believe that my notebook is green, I have a belief about my notebook: that it is green. The intentional object of my belief is the notebook and the intentional content is the proposition that my notebook is green. I accept, however, that a

mental state can be intentional without representing its object propositionally.[10] So, when I look at the notebook, my perceptual state is also intentional even if it does not represent the object propositionally like with belief. My perception is of the notebook, representing it as green, and we could say that the intentional content is a green notebook. To say that both occurrent emotions and attitudes are intentional, then, is to say that they are about or directed or targeted at something and represent that thing in some way. They have an intentional object and content.

Following convention, let us call what an emotion is about or directed or targeted at, the particular object of an emotion. As we saw in the examples in section 1.1, the particular object varies quite substantially. Love, for instance, can be about a concrete object or an abstract ideal: I love my parents but I also love Justice. Fear, too, can be about a concrete object or something more abstract such as an imagined but unlikely state of affairs. I fear the figure on the street but I also fear the prospect of ending up in a small backwards Karoo town raising sheep for the rest of my life. Kim in 5A fears an actual figure with a knife, a present concrete object, but he also fears something in the future, his Actuary exam in 5B. We see, then, that 5A and 5B illustrate the variety of particular objects that emotions can take.

While emotions have particular objects, we have not yet picked out their full intentional nature. In examples 6A and 6B, the particular object of Sasha's fear and the particular object of her love are both the rottweiler. What is it that makes an emotion one of love rather than one of fear? Why is my love for my parents the same kind of thing as my love for Justice?

In the emotion literature, emotions are said to have formal objects which emotions of the same type share and which differ across emotion types (Deonna and Teroni 2012, 41). A formal object of a mental state is a medieval notion that was revived by Anthony Kenny in *Action, Emotion and the Will*, first published in 1963.[11] The formal object of an intentional state, as Kenny reintroduced it, is the description which must apply to the material object if it is to be the object of the state it is (Kenny 1976, 189). The material object is what is feared, loved, found offensive or

[10]There is debate in the philosophy of mind about what intentionality is, whether all mental states are intentional, and what intentional content is. I cannot do more here than stipulate what I take intentionality to be. My reading is influenced by Tim Crane, who provides a useful discussion of how intentional representational content need not be propositional. See Crane (1998), but also the section of essays about intentionality in Crane (2014). Crane distinguishes between the intentional object and the intentional content (and also the intentional mode, which I do not discuss here but which undoubtedly influences my view of emotions). He argues that the defining feature of psychological states is that they have intentional objects at which they are directed. Whether or not he is correct in this further claim, it is useful to distinguish between the object and content.

[11]From here on, I refer to the 1976 edition.

surprising, etc., and is what is seen to have the formal object. Nowadays, people talk about the formal object as a property rather than as a description (Tappolet 2005, 229, fn 1). So, the formal object is some property which must apply to the material object in order for the material object to be the particular object of an emotion.[12] This is usually interpreted as the idea that an emotion is about the formal object; that is, the particular object is represented in a way as being [the formal object]. In this way, the formal object features in the intentional content.

One way of thinking about the formal object of an emotion is as an evaluative property (Deonna and Teroni 2012, 41; Teroni 2007, 396). Fear, for instance, is about something as (having the property of being) dangerous, love is about something as (having the property of being) affection-inspiring, anger is about something as (having the property of being) offensive, sadness about something as (having the property of being) a loss, etc.[13] But in order to be dangerous, affection-inspiring, offensive, a loss, etc., that thing must stand in relation to someone or something. A cliff-face is not dangerous to an eagle which can fly, for example, but it is dangerous to a ground-bound human. Similarly, something which is offensive for me may not be offensive for you. And Kim may fear his Actuary exam but not his Dutch exam, because the Actuary exam poses a danger to his Actuary career but his Dutch exam is no real threat to the success of his hobby of learning Dutch. The formal object is thus also relational.

We can identify five important roles which the concept of the formal object plays in our understanding of emotions. I shall discuss these five roles in more detail in Chapter 6.

First, the formal object is what all instances of an emotion type share. Sasha's fear is elicted by the rottweiler as dangerous and Kim's fear is elicited by the Actu-

[12]The particular object of an emotion need not be the material object. For instance, I may be wrong about the source of my anger but there is still nevertheless something offensive which elicits my anger. That is, the particular object of my anger, what my anger is about or directed at, is not the material object which has the formal object and in fact elicits my emotion.

[13]Amongst 'first flavour' appraisal theories in psychology, a similar notion to the formal object is discussed as the 'core relational theme' (see Moors 2014 for the two 'flavours'). Richard Lazarus is one of the pioneers. See Lazarus (1991), but also Lazarus (2001, 64) for a useful summary and table of core relational themes, including:

Anger	A demeaning offence against me and mine.
Fright	An immediate, concrete, and overwhelming physical danger.
Guilt	Having transgressed a moral imperative.
Sadness	Having experienced an irrevocable loss.
Jealousy	Resenting a third party for loss / threat to another's affection / favour.
Love	Desiring / participating in affection, usually but not necessarily reciprocated.

ary exam as dangerous. Both Kim's fear and Sasha's fear are elicited by and about something as dangerous.

Second, an extension of the first role, the formal object is what individuates emotions (Deonna and Teroni 2012, 41)[14] or defines an emotion type (de Sousa 1987, 126). It is because Kim's and Sasha's emotions are elicited by something as dangerous that they are fear and not love or sadness.

Third, because the formal object is what all instances of an emotion type share, it is a property that a particular object must be seen to have in order for the emotion to be intelligible.[15] If a rottweiler is feared, for example, it must be elicited by something experienced as dangerous. If it is not elicited by something experienced as dangerous, then the fear is not intelligible.

Fourth, the formal object determines whether an emotion is correct or incorrect in the way it represents things as being. Only if there is something dangerous, is fear correct.

Fifth, the formal object determines whether an emotion is appropriate. An emotion can be inappropriate even if it is correct, perhaps for moral or prudential reasons. For example, it may be representationally correct to fear a member of an ethnic group because that person really is dangerous, but inappropriate to fear her because of a moral commitment not to be susceptible to stereotypes.

We can now see what the difference between cases 6A and 6B is. While Sasha's emotions have the same *particular* object, the rottweiler, in each example her emotion is elicited by the dog as having a different formal object. In 6A where Sasha fears the rottweiler, she experiences it as dangerous or frightening. In 6B where Sasha loves the rottweiler, she experiences it as valuable and affection-inspiring. Perhaps in the former, it is the dog's sharp teeth and gruff bark which ground why she experiences it as dangerous; in the latter, it is perhaps the dog's gentle nature and a shared childhood which make her experience it as valuable and affection-inspiring.

We can also now answer the question of why occurrent emotions and emotional attitudes are both emotions. Alex's love for his daughter, whether a state like in 4A or an attitude in 4B, involves regarding her as valuable and affection-inspiring. I love my parents even when I am angry with them because I still regard them as

[14]Teroni (2007) argues that formal objects do not individuate between emotions. Even if the formal object does not individuate, the kind of evaluative nature of emotion that the notion captures is still fundamental to our concept of emotion. I shall discuss Teroni's argument in Chapter 6.

[15]As Ronald de Sousa writes, it 'is a property implicitly ascribed by the emotion to its target, focus or propositional object, in virtue of which the emotion can be seen as intelligible' (de Sousa 2014).

valuable in my life, and the person who fears flying regards flying as dangerous, whether or not she is actually in an occurrent state or not.

Because of the involvement of the formal object, emotions, both states and attitudes, are evaluative and Hacker's evocative phrase that attitudes involve a 'persistent colouring of thought, imagination and fantasy' captures how emotions involve seeing things in evaluatively loaded ways. While it is by no means clear to me that agitations are richly intentional rather than simply reactions, they are at the least elicited by states of affairs which potentially instantiate certain evaluative properties, like getting a fright from something dangerous or feeling disgust on sniffing rotting food waste, and are plausibly about these evaluative properties as well.

This finally puts us in a position to be more precise about the affective nature of emotion. I proposed that their valence is linked to what an emotion is about, and we now see that emotions have an intentional nature that is evaluative.[16] So, while an emotion may involve certain bodily manifestations and sensations, the emotional feeling also has a distinct quality directly related to the evaluative nature of the emotion. I shall talk of this aspect of an emotion's phenomenology as the 'intentional quality' of an emotion.[17] While occurrent emotions tend also to have a bodily manifestation, both occurrent emotions and attitudes have this intentional quality. If I love my parents, then I do think of them with a positive feeling of warmth and comfort and the person who is afraid of flying does have negative feelings when thinking about the prospect of flying, even if they are not strictly sensations.

This leaves the final set of examples to examine, examples comparing emotion and mood. Being in a mood is not like being in a persistent emotional state, either a persistent occurrent state like growing nervousness for an upcoming interview, or an emotional attitude, such as love for one's parents. Unlike an emotion, a mood is not tied to a specific particular object. For example, one may be in a fretful mood which has no immediate object but which involves a general sense of worry and worrying. In such a mood, I worry about the threat of ISIS in the Middle East; I worry about forgetting to turn off the water before I go away on holiday; I worry about the email I sent a friend, in good faith, but which now seems in poor taste; I worry about eating the parsnips at the bottom of my fridge before they go off; I worry

[16] While it is widely accepted that emotions have an evaluative *content*, some argue that the content is not evaluative; rather, it is the attitude. See, for example, Deonna and Teroni (2012) developing ideas found in Goldie (2000). I choose to talk about the 'intentional nature' in order to remain neutral on the disagreements, at this stage at least.

[17] I am only focusing on the feeling as it relates to the content because this is pertinent to my interests. However, as we saw with the brief introduction of valence, emotions are also taken to involve desires and motivational tendencies and elements to the way an emotion feels could also be cashed out in those terms.

about the glancing look someone just gave me at the coffee shop. In such a mood, everything is potentially troublesome and the particular object of my worry swaps and changes. Sam in 7A has a particular object to her sadness, the loss of her cat, but in 7B where she is in a gloomy mood, she does not. This distinction between emotions and moods might lead us to say that moods are not emotions (Hacker 2004) or to say that moods are a variation of emotion (Prinz 2004). Nevertheless, both emotions and moods can be agreed to have a formal object but moods lack a consistent particular object.

We therefore see that we can create some order to the things which we refer to with 'emotion' and the individual emotion words. Emotions are intentional and have both a particular and a formal object. The formal object is evaluative and emotions thus have an evaluative intentional nature. In light of this, emotions are not simply bodily sensations but have an affective nature with a particular phenomenology related to being evaluative, what I have called their intentional quality. Finally, the idea of a continuum with agitations on one end and moods on the other allows us to pinpoint similarities the two extremes share with the more typical emotions, even while having substantial differences.[18] Occurrent emotion states, lying in the middle, are the paradigm emotions.

1.3 Conclusion

Our concept of emotion picks out a range of different things: agitated reactions, occurrent emotion states, emotional attitudes, moods. Nevertheless, by examining sets of examples contrasting emotions with non-emotions and variations within and across emotions, I have shown that this range still shares two important features. They have an intentional nature that is evaluative, and they have an affective nature with an intentional quality.

In terms of their intentional nature, emotions have a particular object and a formal object. The particular object is what an individual emotion is about and that at which the emotion is targeted; it varies across individual instances of emotion. The formal object is the property in virtue of which a particular object is the object of a specific type of emotion; it is what individuates between emotion types, whether an occurrent state or attitude. The formal object is evaluative in nature.

[18] My focus has necessarily been limited on two features that are shared across attitudes and occurrent states. I do not mean to say that these are the only key features to emotions. Paradigmatic emotions might include other important features such as: action tendencies, bodily change, valence cashed out in different terms.

Moods, like emotions, share an evaluative nature with a formal object but are unlike emotions in not having particular objects at which they are targeted.

In terms of their affective nature, while the paradigmatic occurrent emotions involve bodily changes, the way an emotion feels is more general than the feelings of those bodily changes. It includes the phenomenology of being in a particular evaluative state or of having a certain attitude which is evaluative, what may be called its 'intentional quality'. The affective nature is thus closely linked to the intentional nature.

In spite of these common features, it may still be the case that 'emotion' is not a natural kind or even a unitary concept. However, it does not follow that the concept of emotion is confused or incoherent. In fact, it is a concept of a diverse range of things which share important features. It is a rich concept.

But just like an abstract object will have different roles from those of a physical object, despite both being 'objects', the different forms of emotions could have quite different roles in action and explanations of action. Given that we refer to such a range of things as emotions, explaining an action in terms of an emotion could be explaining it in terms of an agitation, an occurrent state, an attitude, a mood, or in terms of how the emotion manifests bodily, what it is about, or in terms of its intentional quality. Perhaps some of these explanations will not be rationalising explanations. But perhaps others of them will. In the following chapter, I present examples of different ways in which we explain action with reference to an emotion and, in Chapter 3, I argue that there is conceptual room for an emotion to provide the agent's reason for acting, and for an explanation in terms of an emotion to be one which explains the action as being done in light of a reason.

CHAPTER 2

EMOTIONS IN EXPLANATIONS OF ACTION

Emotion is often involved in action. Sometimes emotions are at odds with our perception of ourselves as acting rationally: they can warp our judgement, not always respond to our considered judgements, even lead us to act against our better judgements. But sometimes emotions are fully compatible with our perception of ourselves as acting rationally: they can be aligned with or bolster our judgement, motivate us to act timeously, and they can even give us insight into our real values or information about how things are around us. An explanation of action in terms of an emotion could pick up on any of these and we can refer to emotions to explain action that the agent intentionally performs, as well as many instances of behaviour where the subject is not acting in her capacity as an agent.

As we saw in the previous chapter, our concept of emotion is of something rich and refers to a diversity of different things—from agitations, occurrent emotion states and emotional attitudes through to moods, and from bodily sensations through to intentional states. An emotion under different guises could be involved in the different actions and action explanations.

In this chapter, I introduce a range of examples of emotions in explanations of action. I have divided them into five categorises based on what kind of role the emotion plays in the explanation, starting with those where the role of the emotion is tangential to the action *qua* rational and increasing the emotion's involvement with the action *qua* rational, ending with examples where the emotion is crucial to providing the agent's reason—or so I shall ultimately argue. In Chapter 5, I say more about what it means for the agent to have a reason and what rationalising an action involves. For the purposes of this current chapter, it is enough to just point out that paradigmatic rational action involves the agent acting for a reason and in order to achieve something. A rationalising explanation is one which shows that the agent acted or took herself to be acting in light of a reason.

2.1 EXAMPLES OF EMOTIONS IN EXPLANATIONS OF ACTION

I have divided the examples into five categories, A through to E. The categories are divided up in this way to illustrate the variety of actions which can be explained by an emotion, to identify patterns that are recurring, and to have to hand distinctions which can help us contrast and evaluate the explanatory potential of emotion. They need not be mutually exclusive; an action could fall under one or many, or the categories could even be reducible to only one. There may also be more categories. Nevertheless, there is a *prima facie* variety of roles for emotion in an explanation.[1]

CATEGORY A. THE EMOTION AS A MANIFESTATION NOT LINKED TO ACTION

In this category, the emotion provides an explanation because of the way it is manifested. The way it is manifested is independent of whether the agent is acting for a reason, or acting at all, and the explanation is not a rationalising one.

Categories A0 and A1 involve emotion in the same way. Emotions manifest themselves in different ways which can influence how we behave and act, where not all behaviour is action. Even when there is action, some manifestations of emotion do not contribute to an action as rational other than by affecting the way in which the action is performed, or by making it easier or harder to perform the chosen action. Agitations, as defined in Chapter 1, affect action in this way.

Category A0 is not a case of action or behaviour but it sets up a contrast with A1. Both involve physical manifestations of fear but in A0 the manifestation does not affect the agent's chosen action and the action is not what is being explained by the emotion. In A1, it does affect the chosen action. However, in the latter, the manifestation is ultimately irrelevant for explaining the action as rational. Category A2 illustrates how emotions can influence the manner of acting. Again, however, the way the emotion manifests is tangential to understanding the action as rational.

[1] In the examples that follow, I draw loosely on characters from a range of sources. Hawkeye, McIntyre and Frank are from *M*A*S*H*. Sissy Jupe is from Charles Dickens, *Hard Times* (first published in 1854, using 1905 edition). Huck Finn and Jim are from Mark Twain, *The Adventures of Huckleberry Finn* (1885). Alice is from Patricia Greenspan, 'Craving the Right: Emotions and moral reasons' (2011). Begbie is from Irvine Welsh, *Trainspotting* (1993). Natasha, Pierre, Rostov, Dolokhov and Prince Andrey are from Leo Tolstoy, *War and Peace* (first published in English in 1886, using 2001 edition). Othello and Desdemona are from William Shakespeare, *Othello* (first performed around 1604, using 1997 edition). Emily is from Nomi Arpaly, 'On Acting Rationally Against One's Best Judgment' (2000). Claire is discussed on the NHS website (NHS 2011). Where I stick directly to the source and do not manipulate or interpret the characters, I give the direct reference.

A0. WHERE THE MANIFESTATION OF THE EMOTION IS 'INERT'

- *Hawkeye sweats because he is afraid.* —While Hawkeye performs intricate surgery on a wounded soldier, he sweats from fear. He is a professional and competent surgeon, however, and, other than a damp brow, his fear is not evident. It does not affect his carefully executed movements.

A1. WHERE THE MANIFESTATION AFFECTS THE EXECUTION OF AN ACTION

- *Hawkeye drops the spanner because of his fear / out of fear.* —While Hawkeye defuses an unexploded bomb in the middle of the campsite, he sweats from fear. Out of his comfort zone, his hands also start shaking.

A2. WHERE THE EMOTION IS MANIFESTED IN PARTICULAR MANNERS OF BEHAVIOUR

- *Frank has a sharp manner because he is angry.* —When Frank gets angry with Hawkeye and McIntyre, he stamps his feet when he walks and slams his food tray down on the table with more force than if he were not angry. He speaks louder and snaps brief answers. While in surgery, his anger does not affect his choices or his skill and the only way you would know he was angry is from his sharply toned voice and snappy answers.

- *Sissy Jupe cowers out of fear.* —Picked on by Thomas Gradgrind to define a horse, Sissy Jupe is overwhelmed by fear. She blushes, curtsies and cowers out of fear.

- *Kim jumps out of fright.* —Rounding a corner late at night, Kim gets a fright when he comes face-to-face with a figure holding a knife-shaped item.

CATEGORY B. THE EMOTION AS EXPLANATION OF, OR ALLUDING TO, OTHER MENTAL STATES

In this category, the emotion provides a context where the agent has certain beliefs, desires or attitudes. As I argue in Chapter 4, emotions are not constituted by beliefs or judgements. Nevertheless, emotions can and do lead to associated beliefs and desires, lead us to judge in certain ways, or influence the kinds of beliefs and desires that we form. This is because of the way they have intentional content, as I argue in Chapter 3. In the examples here, the emotion is an attitude or an occurrent emotion state, but it is the subsequent mental states—or, strictly speaking, the contents of those mental states—which would play a rationalising role in the explanation of the

action. The emotion is informative in explaining why the agent has those mental states in the first place, or in explaining why the agent has the attitude towards her surroundings that she does.

B1. WHERE THE EMOTION IS CLOSELY LINKED TO OTHER BELIEFS AND DESIRES, PERHAPS BY CAUSING THE BELIEFS AND DESIRES

- *Huck Finn failed to turn Jim over to the authorities because of his compassion for Jim.* —Despite believing that he should turn the escaped slave Jim over to the authorities, Huck Finn fails to do so because he feels compassion for Jim. His compassion leads Huck to believe that Jim is a good and dear human being and to desire to protect him. When we say that Huck acted out of compassion, we are alluding to these further beliefs and desires to make sense of why he did what he did.

- *Alice challenged her boss out of moral indignation.* —Alice challenges her boss' decision to block a deserved pay rise for a colleague because she feels moral indignation. She finds herself feeling indignant and comes to believe that she ought to challenge the injustice.

B2. WHERE THE EMOTION AFFECTS HOW WE SEE A SITUATION, IN TURN AFFECTING WHAT ACTION THE AGENT PERFORMS OR HOW THE AGENT BEHAVES

- *Begbie beats up the man because he is angry.* —Begbie is losing at pool and getting increasingly angry. After a particularly bad shot, he beats up a young man at the bar because, he says, he was put off by the man's looking at him. We explain his action as being done because he is angry, where the anger leads Begbie to see the man as offensive regardless of whether the man is offensive or not.

- *Frank kicks up a fuss because he is angry with Hawkeye and McIntyre.* — Frank finds everything that Hawkeye and McIntyre do offensive and believes that they are out to get him. This makes him quite angry. It just so happens that they are really out to get him most of the time; nevertheless, Frank cannot tell between cases where the two are really doing something just to spite him or where they have not done anything wrong at all. Regardless, his anger makes him extra-sensitive to everything that Hawkeye and McIntyre do and he sees all of their actions as offensive.

- *Pierre gives alms to the poor because of his overwhelming love.* —Married to Natasha, Pierre's 'heart was now overflowing with love, and by loving people without cause he discovered indubitable causes for loving them' (Tolstoy 2001, 971). Pierre does not need to look for good qualities in people; his love makes him see good qualities everywhere. Because of this, he chooses to do things that benefit others.

B3. Where the emotion is the agent's motive for acting

- *Frank reports Hawkeye and McIntyre to the commander out of anger.* — Thoroughly fed up with being picked on, Frank reports the troublemakers to the captain. A bit later when he has cooled down and is asked why he reported Hawkeye and McIntyre, he responds that he did so out of anger and he would do so again because they deserve to be punished. Frank acts out of the motive of anger even if he need not be in an occurrent angry state at the time.

- *Othello kills Desdemona out of jealousy.* —Othello comes to believe that his wife Desdemona has been unfaithful. He sees her guilt in her proclamations of innocence and kills her out of the motive of jealousy. We could also say that he killed her out of love, as Othello himself identifies. He loved her 'not wisely but too well' (Shakespeare *Othello*, Act V scene 2) and could not handle losing her to another.

Category C. The emotion as that which is expressed in purely expressive action

Unlike the examples in Category B, the next three categories provide examples of cases where the relevant state playing a role in the action is the emotion itself.

In this third category, the action is an expression of an occurrent emotion. The examples are specifically of pure expressive action which Rosalind Hursthouse characterises as action where 'I ϕ-ed because I was so frightened (or happy, excited, ashamed . . . so overwhelmed by hatred or affection or . . .) that I just wanted to, or felt I had to' (Hursthouse 1991, 58). Hursthouse identifies three conditions for this kind of action to meet: the action must be intentional; the agent did not do it for a reason 'in the sense that there is a true description of action of the form "X did it (in order to). . . " or "X was trying to. . . "'; and the agent would not have performed the action if she were not in the grip of an emotion (59).

Sometimes, one might express an emotion in order to achieve some goal such as

to communicate one's frustration or release pent-up energy, examples which would fall under the next category. But the case of interest here is where we are just expressing the emotional state we are in for no purpose other than that we can and we want to. It is controversial whether the agent is acting for a reason at all in these cases. I will address expressive action in Chapter 5.

- *In a wave of hatred, Jane tears at Joan's eyes in a photograph* (Hursthouse 1991, 59). —Jane is not trying to communicate her anger nor relieve her pent-up energy. She acts purely because she is angry.

- *Having lost his much-loved wife, the man rolls around and buries his face in her clothes* (Hursthouse 1991, 58). —Like Jane, he is not communicating his grief or relieving his emotion; he is simply expressing it.

CATEGORY D. THE EMOTION PROVIDES THE AGENT'S REASON AS A STATE TO BE ADDRESSED

The last two categories I shall introduce examples of different ways in which an emotion could be or provide the agent's reason and thus potentially play a rationalising role in an explanation. The examples I describe are primarily of occurrent emotion states (except Dolokhov's mood in example D1(a) and Natasha's attitude in E) and in discussing these categories in the following chapter, I will focus on occurrent states. In this category, D, the emotional experience is the agent's reason for acting in order to avoid, alleviate or encourage the emotional experience in some way.

D1(A). ADDRESSING THE SYMPTOMS OF THE FEELING OR MANIFESTATION OF THE EMOTION

- *The violinist took a sedative because she was nervous.* —When I was playing in an amateur orchestra, the conductor used to ask people who could not control their nerves during performances to take a light-weight sedative beforehand. These people took the medication in order to alleviate the symptoms of their nervous fear, such as shaking.

- *Begbie picks fights because he enjoys being angry.* —Begbie enjoys getting angry and beating people up. One evening while sitting upstairs at the pub, he throws a glass over the bannister which hits a woman on the head. He then goes downstairs and picks a fight in order to get the person who harmed the

woman to own up, knowing full-well that he is the culprit. Invigorated by his anger, he seeks ways to encourage it.

D1(B). ADDRESSING THE WAY THE EMOTION IS MANIFESTED IN BEHAVIOUR

- *Claire avoids open spaces out of fear.* —Claire has been diagnosed with agoraphobia. She is afraid of open spaces. After a panic attack in her local high street, she believed that the location was behind her inexplicable experience and began shopping elsewhere. However, she had similar attacks in other locations. In order to avoid the panic attacks, she chose to leave her job and spend all her time indoors.

- *Dolokhov challenges Rostov to gamble away money he does not have out of cruelty.* —'Beneath his smile Rostov saw in him [Dolokhov] the mood he had shown at the club dinner and at other times, when as tired of everyday life he had felt a need to escape from it by some strange, and usually cruel, action' (Tolstoy 2001, 287). Dolokhov suffers from apathetic and depressive moods. These are objectless and he attempts to alleviate them by performing strange and cruel actions.

There is nothing particularly special to emotion in these two cases. For instance, in the first example in D1(a), what is being addressed is the shaking. Because it is shaking brought on by fear and not, say, a nervous disorder, a light-weight sedative is the right kind of thing to deal with it. The emotion causes the shaking, and the kind of cause thus opens up what kind of things could allow the person to address it, much like the cause of a sore throat will determine what treatment is best. In the first example in D1(b), we can imagine an analogous scenario where Claire is violently sick after eating mussels. She chooses to avoid the restaurant where she had them, believing that the location was behind her sudden sickness. On eating mussels elsewhere and becoming sick again, she now might choose to avoid mussels altogether. In both the agoraphobia case and the mussel case, the emotion and the sickness are unpleasant negative states which the agent acts to address. One just so happens to be an emotion.

Category D2, in contrast, involves the emotion as more than a feeling or manifestation. The agent acts because of the emotional experience as a whole, including the intentional element.

D2. NOT JUST ADDRESSING THE FEELING OR MANIFESTATION

- *Jackson throws out his ex-partner's belongings out of grief.* —Struggling with a horrible break-up and feeling debilitating grief over his loss, Jackson eventually goes to see a counsellor. The counsellor recommends that he collect all of his former partner's belongings and take them to the junk shop. At first, Jackson is dubious and begins only because the counsellor told him to do so. However, he finds the whole process incredibly cathartic. As he continues to gather the items, he no longer does so only because the counsellor told him what to do but also because he gets relief from his anguish. He only gets that relief because they are his former partner's belongings and he associates them with his loss. He does not get the same relief from helping his parents with their spring-cleaning.

- *Alice challenges her boss because she is indignant.* —Alice is indignant about her boss's treatment of her colleague. She is all a-flutter, which leads her to act spontaneously and immediately in challenging her boss.

CATEGORY E. THE EMOTION PROVIDES THE AGENT'S REASON, BUT IS NOT THAT REASON ITSELF BY BEING A STATE TO BE ADDRESSED

In this last category, E, the agent acts because of the emotion but not in order to address a positive or negative state or to express it. Rather, she acts in light of a reason provided by the intentional nature of an emotion.

- *Richard bought flowers for Katherine because he loved her.* —Buying flowers for a loved one need not be an expression of caring for that person, nor a means to encourage the love. Rather, buying flowers is part of actually caring for and loving that person.

- *Natasha pulls her life together out of love for her mother.* —Natasha is distraught over the death of Prince Andrey. She becomes ill and loses all purpose to her life. However, her love for her mother 'unexpectedly showed her that the essence of life—love—was still active within her' (Tolstoy 2001, 933). Through her love for her mother, Natasha starts to pull her life together. She is neither avoiding or encouraging the state, nor expressing it. Yet, it is true to say that she acts as she does *because she loves her mother*, where the love explains more than that she has a particular attitude, as Pierre's love does in example B2.

- *Huck Finn does not turn Jim over to the authorities because he feels compassion for Jim.* —Unlike in B1, Huck Finn does not form other beliefs about Jim or what he ought to do, he simply acts out of his compassion.

- *Coming across a rhinkals in the path, John backs away out of fear.* —John does not jump backwards out of fright like Kim in A2, but he also does not deliberate about what to do. He backs away slowly.

- *Alice challenges her boss because she is morally indignant at the injustice.* — Unlike in D2 where Alice challenges her boss because she is all a-flutter and acts spontaneously, Alice challenges her boss because of what she is morally indignant about.

- *Emily quits the PhD programme because she is sad and ill-motivated.* — Emily, a PhD student in chemistry, finds herself feeling sad and ill-motivated in her endeavours, despite having always believed that she should pursue a PhD. She sees the feelings that make her want to leave the programme as groundless even though, unknown to her, they are triggered by factors that would be good reasons not to be doing the PhD. She eventually leaves the programme on an emotional impulse and judges herself to be acting irrationally. When she looks back years later, however, she views her persistence as irrational and takes as her reasons for acting the reasons behind her emotional state (Arpaly 2000, 504).

The following table summarises the examples:

A	Manifestation	A1. Symptoms of manifestation A2. Manifestation in behaviour
B	Other intentional states	B1. Beliefs and desires B2. 'Seeing' B3. Motives
C	Pure expression	C. Expressing an emotion
D	As a reason: addressing a state	D1. Avoiding, alleviating, encouraging: (a) Symptoms of manifestation (b) Manifestation in behaviour D2. Avoiding, alleviating, encouraging the emotion, not just the manifestation
E	Providing a reason: intentional nature	E. Not avoiding, etc., or expressing

2.2 CONCLUSION

The examples in this chapter illustrate the ways in which we can give an explanation with reference to an emotion. I shall refine the examples to illustrate what role, exactly, the emotion plays in Chapter 3 (Categories A, B and D), Chapter 5 (Category C), and Chapters 6 and 7 (Category E). For now, however, the examples are enough to illustrate the variety of roles. Drawing this together with the first claim, that our emotion concept is rich and diverse, we now see that, even if an emotion does not play a rationalising role in one kind of explanation, it need not follow that it does not play a rationalising role in another kind. This is because an emotion could play a role in virtue of being an agitation, an occurrent state, an attitude or a mood, or in virtue of how it feels, what it is about, or both together. And, moreover, if it plays one kind of role in one kind of explanation, it need not play the same role in all explanations.

Despite this, there has been a tendency in the history of philosophy to treat emotions as having only one or at most two kinds of role in action, and a tendency to think that their role in an explanation will not be a rationalising one. I shall examine this tradition in the next chapter. I argue that developments in the twentieth century which highlight the intentional nature of emotion as well as distinguish between the agent's reason for acting and the cause of the action are steps in the right direction, but that our conception of the potential explanatory role of emotion can be enriched further following increased understanding of what emotions are.

CHAPTER 3

HISTORICAL OVERVIEW OF EMOTIONS AND ACTION

We have seen how rich and diverse our concept of 'emotion' is, and how different the ways emotions can potentially explain action are. In this chapter, I give a brief overview of various historical themes to the treatment of emotions and action, and argue that historical treatments of emotion in action and explanations of action involving emotion have been too heavy-handed.

I begin in section 3.1 by identifying three general trends that have dominated pre-twentieth-century thought about emotions and action.[1] These are: (i) the emotions, as passions, were seen to be passive and opposed to activity, like perceptions; (ii) the emotions were seen as states that straddle both the mind and body; and (iii) the emotions were given a motivating role in action. While emotions have often been opposed to reason or said to operate tangentially to reason, I shall suggest that these themes taken abstractly from their historical contexts need not be at odds with the pro-emotion consensus, the school of contemporary views endorsing the claim that emotions are compatible with reason. In particular, if emotions are passive in a way like perception, can emotions not play a role in action much like that of perception? If emotions are not just states of the body and involve some kind of thought or perception which has content, can they play a role beyond simply motivating action?

In section 3.2, I turn to some twentieth century developments and focus primarily on the accounts of emotion and action put forward by Gilbert Ryle and Anthony Kenny. Using the categories of examples introduced in Chapter 2, I argue that Ryle's account of emotion and action is unable to deal with examples like the pure expressive action in Category C. The failure arises because of his inadequate conception of emotion which does not treat emotions as intentional states. Kenny

[1] I am indebted to Susan James' informative book about the emotions in seventeenth-century philosophy and their historic influences, *Passion and Action* (1997).

substantially improves Ryle's account because he introduces the formal object of an emotion and the idea that emotions have targets which they are about. I shall argue, however, that his account also fails with regards to Category C, and that it is too general, thereby remaining vague on important differences between the explanations in categories B and D. This is despite his important insight that emotions are intentional. I shall suggest that, if we push Kenny's conclusions a little further, a whole range of other options open up for emotion in action and explanations of action. I therefore end the chapter by concluding that there is conceptual space for emotions to have a reason-providing role in action and in explanations of action.

3.1 HISTORICAL OVERVIEW

> For we see reason is disturbed ... by passion or affection ... And as in negotiation with others, men are wrought by cunning, by importunity, and by vehemency; so in this negotiation within ourselves, men are undermined by inconsequences, solicited and importuned by impressions or observations, and transported by passions. Neither is the nature of man so unfortunately built, as that those powers and arts should have force to disturb reason, and not to establish and advance it. For ... the end of morality is to procure the affections to obey reason, and not to invade it. (Bacon 1893, 68)

> Though this evil [slavery] be so tedious, and the good [freedom] it deprives us so pleasing, it is not comparable to that which the tyranny of the Passions causeth in us. And it must be granted, that of as many slaves as are in the world, there is none more unhappy, than those who obey such cruel Masters. (Senault 1649, 95)

> If we examine carefully why people usually adopt one view rather than another, we will find that it is not the penetration of truth or the force of reason, but some connection to self-love, interest, or passion. This is what carries weight and what causes most of our doubts. It both unsettles our judgements and fixes us most firmly to them. We judge things not by what they are in themselves, but by what they are in relation to us. For us truth and utility are the same thing. (Arnauld and Nicole 1996, 204)

We can see from these quotations the way in which emotions have historically been opposed to reason: emotions disturb reason; they need to be constrained; they affect how we see the world but are not themselves necessarily connected to the truth.

Emotions have therefore not always been viewed favourably in the way that those forming the pro-emotion consensus in contemporary philosophy of mind and psychology view them: not in conflict with reason, and as fundamental to our capacity to operate as rational agents and even as providing us with information about how things are.

In this short first section, I draw attention to three themes that have been central in historic understandings of the emotions and propose that we can find variations of them in contemporary discussions of emotions. In doing so, I aim to show how we need not be forced into the negative views expressed in the quotations above, where emotion acts tangentially, at best, or contrary, at worse, to reason.

The first theme is that emotions, understood as passions, are passive and opposed to activity, like perceptions. As Susan James (1997, 30) notes in her comprehensive discussion of emotions in seventeenth-century philosophy, Aristotle's notions of activity and passivity were influential amongst seventeenth-century philosophers. Aristotle holds that the emotions, like other passions,[2] are passive. This is because emotions, for Aristotle and his followers, were seen as not having active powers to effect transformations. They were rather seen as a capacity to be affected by things around us (41).

James favours interpreting Aristotle as saying that emotions are similar to sensory perception. Like sensory perception, emotions were seen as 'responses that have to be invoked in us by external things and as states that we suffer. We do not have the power to experience passions unaided, but must wait on circumstances to excite them' (James 1997, 42). The Aristotelian idea of emotions as passive in some way gets picked up by others, such as by St Thomas Aquinas in the *Summa Theologiæ* Volume XIX and by René Descartes, who writes in *The Passions of the Soul* that the passions are 'perceptions, sensations or emotions of the soul' (Article 27). Emotions, on this understanding, are things that happen to us.

The second theme is that emotions were seen as straddling both the mind and the body. Aristotle, for example, writes that emotions are:

> . . . those things through which, by undergoing change, people come to differ in their judgements and which are accompanied by pain and pleasure, for example, anger, pity, fear, and other such things and their opposites. (Aristotle, *On Rhetoric*, 1378a 20-23)

[2]The use of 'passions' is not co-extensive with contemporary use of 'passions' or of 'emotions'. For one thing, Aristotle includes appetite amongst the passions, and we often find desire included as well (James 1997, 40-41).

In *On the Soul*, we find:

> Do the conditions of the soul belong as well to the body or is there one that belongs only to the soul itself? We encounter this difficulty here, and it is a difficulty neither avoidable nor easily treated. It seems that most of the soul's conditions—anger, courage, desire, and any sensation—neither act nor are activated without the body. (Aristotle, *On the Soul*, 403a 2)

Aquinas argues that, while emotions are seated in the soul, they are seated in the sensory or orectic aspect. This is because, for Aquinas, the passions involve being drawn to something good or bad; something is good or bad in itself; and being drawn to things as they are in themselves requires physical modification. He writes:

> The term 'passion' implies, as we have said, that the patient is drawn to something in the agent. But the soul is drawn to things by its orectic, rather than its cognitive, faculties. For through its orectic faculties the soul is drawn towards things as they are in themselves: as Aristotle says, good or bad (the objects of the orectic faculties) are in themselves. (Aquinas, *Summa Theologiæ*, 22.3)

Later, Descartes distinguishes between the body and the soul, placing the emotions in the soul but emphasising that they are influenced by the 'animal spirits' of the body. He writes:

> . . . the passions of the soul are perceptions, sentiments, and emotions of the soul, which are referred particularly to the soul itself, and which are caused, entertained, and strengthened by some movement of the animal spirits. (Descartes, *The Passions of the Soul*, Article 27)

Baruch Spinoza, unlike Descartes, identifies the mind with the body and argues that the passions function to promote our whole self (James 1997, 147). They are responses to the world around us, as it affects our whole being, body and soul.

These are just some examples of how emotions have been understood as straddling the mind and the body.

The third key theme is that emotions have a motivating role in action. Aristotle, for instance, defines an emotion like anger as a desire for vengeance, where a desire motivates us to act. For Aquinas, while emotions are passive, they still involve movement towards the good and, ideally, we must align our emotions with our conceptions of the good. For Descartes, the passions involve 'some movement of the animal spirits' (*The Passions of the Soul*, Article 27) and:

> . . . it is requisite to notice that the principle effect of all the passions in men is that they incite and dispose their soul to desire those things for which they prepare their body, so that the feeling of fear incites it to desire to fly, that of courage to desire to fight, and so on. (Descartes, *The Passions of the Soul*, Article 40)

Descartes keeps perceptions and volitions distinct, but Spinoza argues that they are part of the same thing, thereby allowing that being moved to act can be part of a perception such as an emotion. Finally, as we saw in the quotations at the beginning of this section, emotions were often seen to have a disruptive effect on reason and judgement. We can interpret philosophers who endorse such views as acknowledging the motivating role of emotions, else the emotions would not cause us to do what we do, but denying that the emotions have a further connection to reasons. For instance, Antoine Arnauld and Pierre Nicole in *Logic or the Art of Thinking*, first published anonymously in 1662, argue that taking our passions and interests as a motive to act or believe is a fallacy common in everyday life, where a motive, for them, is used synonymously with a reason.[3] They write:

> The most [our passion or interest] can do is make us pay more attention to the reasons that could reveal the truth to us about what we want to be true. But we should be persuaded only by this truth, which must be found in the thing itself independently of our desires. (Arnauld and Nicole 1996, 204)

I wish to draw attention to these three themes because of the way in which, abstracted from their historical contexts, they show up in contemporary discussions of emotion. First, emotions are analysed in similar terms to sensory perception even today. As I shall discuss in Chapter 4, a popular contemporary philosophical account of emotion models emotion after perception. The important idea that is explored in contemporary discussions but not so much in historical ones is that, if emotions are passive in a way like perception, then can we gain information about how things are through emotions? If we can be 'acted on' by the things around us then, through our experiences, we could plausibly form beliefs and knowledge about what is acting on us.

Second, emotions involve both the body and the mind. As I have already mentioned in Chapter 1 and shall discuss further in Chapter 4, our concept of emotion is fundamentally of something which is intentional but which also has a distinct feeling which is often a bodily sensation. If emotions are partly of the mind and have

[3]Contemporary usage of 'motive' is of something distinct to a reason. I shall say more about motives in the following section.

an intentional nature, then they could perhaps play a role in action in virtue of what they are about rather than simply how they feel in the body.

Finally, emotions may be intrinsically motivating, perhaps like volition or desire. However, if we take the first two features seriously—that emotions are like perception and can give us information about how things are, and that emotions have an intentional nature—then there could be more to emotions than simply as motivators. There is conceptual space for our acting in light of what emotions are about and not only because of how they feel.

3.2 TWENTIETH-CENTURY DEVELOPMENTS

The ideas which extend the three historical themes and with which I closed the previous section begin to be explored in the twentieth century. These were the ideas that: (a) if emotions are like perception, they could give us information about how things are; (b) emotions have an intentional nature and are both in body and mind; and (c) there is conceptual space for them to be involved in action more than just as motivating states. Two important developments in particular help to open up a new understanding of emotion and action and both relate to causes.

The first distinction is between the cause of an emotion and the object of an emotion, a distinction which Ludwig Wittgenstein (1958), Anthony Kenny (1976) and Ronald de Sousa (1987) influentially make, but which is by no means limited to them. We find Wittgenstein in the *Philosophical Investigations* writing:

> We should distinguish between the object of fear and the cause of fear. Thus a face which inspires fear or delight (the object of fear or delight), is not on that account its cause, but—one might say—its target. (Wittgenstein 1958, §475)

By making this distinction, Wittgenstein opens up for discussion how emotions are intentional states that have targets at which they are directed, but where the targets need not be their cause. In being targeted at something, emotions are mental states with direction and which are about their target. This allows that emotions could be caused by one thing—perhaps we are passively acted on—but they could nevertheless be targeted at or about something else.

The second is a distinction between the cause of an action and the reason for or intention behind the action, a distinction which both Wittgenstein and Elizabeth Anscombe (1957) make.

In conversation with Friedrich Waismann, Wittgenstein, for example, distinguishes between the cause of an action and a reason:

> Let us suppose a train driver sees a red signal flashing and brings the train to a stop. In response to the question: 'Why did you stop?' he answers perhaps: 'Because the signal says to stop here'. One wrongly regards this statement as the statement of a *cause* whereas it is the statement of a *reason*. The cause may have been that he was long accustomed to reacting to the red signal in such-and-such a way or that in his nervous system permanent connections of pathways developed such that the action follows the stimulus in the manner of a reflex or yet something else. The cause need not be known to him. By contrast, the reason is what he states it is. (Wittgenstein and Waismann 2003, 111)

Anscombe also distinguishes between the cause of an action and the reason, motive or intention for which it is done. For instance, if the answer to the question 'Why did you kill him?' is 'He killed my father', Anscombe writes, this 'is surely a reason rather than a cause' (Anscombe 1957, 10) and goes on to elaborate the difference between a mental cause, a reason, motive and an intention. Donald Davidson later opposes the distinction made by Wittgenstein and Anscombe and their followers, and argues that reasons *are* causes (see, in particular, Davidson 1963). Much of recent philosophy of action has focused on whether reasons are causes or not, and the question is still not conclusively settled.[4]

This distinction between cause and reason is important in two ways. First, if the cause of an action need not be the same as the agent's reason for that action, then we can accept that emotions cause us to act but allow that they could also provide the agent's reason in some way. So, even if emotions are motivating states which sometimes cause us to act against our reasons, it does not follow that this is the only role emotions could play in action or action explanations. Second, if emotions do have a motivating role in action, which they surely do, we can ask whether that motivating role is in virtue of causing the action or in virtue of providing the agent's reason in some way, perhaps by being the reason or providing access to the reason. Again, even if it is sometimes the former, it is possible that it can also be the latter.

By distinguishing between the cause and the reason for an action, we immediately allow at least two ways an emotion could be involved with action. The emotion could cause an action or it could otherwise be involved in providing the agent with a reason. One way in which it could be involved in providing the agent

[4]See section 3 in Wilson and Shpall (2012) for some of the key positions and arguments, such as distinguishing between causal and teleological reasons.

with a reason is by being a motive, where a motive is not a reason but is still distinct from the cause. In the rest of section 3.2, I shall address two versions of accounts specifically addressing emotions which distinguish between the cause of an action and an agent's motives, those of Ryle and Kenny who both argue that emotions function either as causes or as motives in an explanation. If they are right, then the only roles available to emotion in action explanations will be like those in Category A, as cause, or Category B, as motive. I argue that Ryle's account fails because of his limited conception of emotion. Kenny's account, however, starts to take us in the right direction because he introduces the formal object of an emotion as what an emotion is about and this opens up the possibility of a connection between emotions and reasons. I shall argue, however, that the idea can take us further than Kenny allows. This is because he gives a too-general pattern for acting from a motive, thereby obscuring ways in which we could also act for a reason when acting from an emotion.

3.2.1 Ryle on causes and motives

In *The Concept of Mind*, first published in 1949, Ryle writes:

> There are two quite different senses of 'emotion', in which we explain people's behaviour by reference to emotions. In the first sense we are referring to the motives or inclinations from which more or less intelligent actions are done. In the second sense we are referring to moods, including the agitations or perturbations of which some aimless movements are signs. (Ryle 1990, 110)

As we can see in this quotation, Ryle identifies two roles for emotion in explanations of behaviour and action, one as motives and the other as moods with causal powers. He even identifies a third sense of emotion, the 'pangs and twinges that are feelings of emotion', but denies that this sense has any fundamental role in an explanation of behaviour. These emotions 'are not, save *per accidens*, things by reference to which we explain behaviour' (Ryle 1990, 110). All explanations of action with reference to emotion will thus be in terms of either motives or as causes.

Ryle uses his own terminology which needs explicating. For him, 'emotion' refers to motives (or inclinations), moods, agitations, or feelings. Motives are dispositions of the person to act. Moods are temporary conditions that 'collect occurrences' (Ryle 1990, 81); being in a sad mood, for example, means that one is liable to experience various occurrences over a length of time, such as to have certain feelings, or to have certain motives or agitations. Agitations (or commotions)

are propensities or liabilities to undergo a disturbance resulting in symptoms such as cowering or shivering or sweating (92-94); and feelings are 'the sorts of things which people often describe as thrills, twinges, pangs, throbs, wrenches, itches, prickings, chills, glows, loads, qualms, hankerings, curdlings, sinkings, tensions, gnawings and shocks' (81). Ryle argues that only feelings are occurrent states of the agent; the rest are propensities.

For Ryle, an agitation is a liability to undergo a disturbance, and is the conjunction of a general inclination or propensity to behave in a certain way and an inhibition of some kind upon such behaviour. What results is expressive behaviour, or signs of the agitation. In this way, the emotion provides a causal explanation: a general inclination to do x together with an inhibition to do x causes the subject to do y or exhibit certain manifestations z. The examples in Category A illustrate this idea well. Frank's anger makes him snap answers in a sharply toned voice. His angry mood is a collection of agitations: he is agitated and disposed to undergo a disturbance, such as to lash out. He is, however, inhibited from doing so, and the result is his sharp tone of voice, a sign of the disturbance. Hawkeye's sweat and shaky hands are signs of his agitation of fear as he defuses a bomb, as is Sissy Jupe's cowering in the face of the terrifying Gradgrind. The emotions in Category A thus fall under Ryle's second sense of emotion as mood or a collection of occurrences, including agitations. The emotion in these kinds of explanations function as a cause that is not the agent's reason and does not explain an action *qua* rational.

Unlike agitations, explaining action with reference to a motive is showing that the behaviour, or action, was performed for a purpose or a reason, and so deals with 'more or less intelligent actions'. Ryle argues that an explanation making use of inclinations or motives is one which draws on tendencies or propensities of the agent to behave in certain ways; these are dispositional properties (Ryle 1990, 83). He writes that '[t]o say that he did something from that motive is to say that this action, done in its particular circumstances was just the sort of thing that that was an inclination to do' (90). Together with a general disposition, we have the circumstances at a particular time that led to the action. In such an explanation, we would be explaining with reference to an emotion by picking out both the disposition and the emotion-context in which the disposition is actualised. Saying that someone acted out of fear is thus saying that this action, done in these circumstances, is just the sort of thing she was inclined to do.[5]

[5]Anscombe objects that, on Ryle's analysis, we can not act from a motive only once but this is blatantly false (Anscombe 1957, 20-21).

We can extend Ryle's dispositional analysis to beliefs and attitudes. In this way, we can see how the explanations in Category B could fall under Ryle's motives: in this situation, having this motive, forming these beliefs or seeing things in this kind of way is just the kind of thing the agent is inclined to do. In categories D and E where the emotion is the agent's reason, we could similarly make sense of the examples as picking out inclinations of the agent to behave in certain ways.

Category C, containing examples where the agent acts purely expressively, however, is not compatible with either causal or motive explanations, and looking at it highlights the insufficiency of Ryle's account of the nature of emotion explanations *tout court*.

Category C contains examples of pure expressive action, such as Jane's tearing at the eyes in a photograph out of rage and the man's rolling around in his dead wife's clothes out of grief. For Ryle, expressive behaviour is the sign of an agitation and the result of being unable to behave in the way to which one is inclined. In the cases in C, however, the person is acting intentionally and doing exactly what she wants: Jane is behaving exactly in the way to which she is inclined. Her tearing at the photograph is not expressive and the sign of an agitation in the same way as Sissy Jupe's cowering. Explanations of pure expressive action thus do not make use of an emotion as an agitation, and so we must see how they fare with the emotion as a motive.

Explaining by reference to a motive, for Ryle, picks out the disposition and the context. In order for the explanation to be satisfactory, there must be some illuminating connection between the particular circumstances of the context and the relevant disposition. The circumstances must be what is occurrent which, crucially, for Ryle are only the feelings. This, however, does not provide a satisfactory explanation.

Consider the example of the man who rolls around in his dead wife's clothes because he is grieving for her. On Ryle's account, when we explain the man's action by citing his grief, we are saying that, in this particular circumstance, rolling around in the clothes is just the sort of thing he was inclined to do. In this particular circumstance that we have identified as grief, he is experiencing the feelings that we label 'grief', perhaps things like having a tight chest and an overwhelming urge to cry, and is disposed to act in a grief-like way, such as to try to be close to the lost loved one. However, there is nothing about the *feelings* of grief that provide a base for the disposition to roll around in the dead wife's clothes. As I shall discuss in more detail in Chapter 4, the feelings of an emotion are not always enough to distinguish one emotion from another, and different emotions with similar or the

same feelings might involve quite different dispositions to act. For example, I may feel the racing heart of anger and be disposed to approach and attack my opponent. I may also feel the racing heart of fear and be disposed to get away as fast as possible. Sometimes I even confuse my fear for anger, as when I think I am angry with my innocuous passenger *just for being there* as I drive through a dangerous part of town. One can also have the same feelings but no disposition to act in a way particular to the emotion-feelings. Why is the grieving man not disposed to roll around in clothes when he has the same tight chest and teary eyes as when he has a frustrating debate at work or when he is chopping onions with a chest cold? What is it about the feelings we refer to with the emotion word which dispose someone to act in a way so typical? The feelings, alone, are not enough to account for the disposition. We have not yet got to the bottom of the role of the emotion in the explanation of the action.

If we can say what it is about emotion feelings that makes them different to other bodily feelings, we can begin to grasp a connection between the feelings and the disposition. And Ryle does distinguish between emotion feelings and bodily feelings. For him, the difference is something we learn through observing our and others' feelings (Ryle 1990, 101-102).

Observing a difference, however, is not enough to provide a satisfactory explanation of expressive action. Merely observing a difference does not pinpoint what it is that we are observing or what it is about feelings in one context that makes them emotion feelings and in another context mere bodily feelings. Neither does it pinpoint what it is about feelings in one context that lead to actions of a certain kind, like fighting, and in another context lead to other actions, like running away. It is not enough to say that in one context, but not another, the feelings are accompanied by a disposition to act in a certain way, because that accompaniment is what we are trying to understand. When we explain by referring to an emotion, even if the emotion picks out a disposition to behave in a certain way, the explanation only gets off the ground because we know what the connection between the context and the disposition is, beyond their co-occurring. We are drawing on more than just how the emotion feels, which is the only thing that Ryle allows as an occurrent emotion.

This criticism of motives does not just show that Ryle's account fails to make sense of expressive action in C, but also that the motive explanation as a whole is ultimately unsatisfactory. The key problem with Ryle's account is that he identifies the occurrent state as only a feeling. As I noted in Chapter 1, and shall develop further in Chapter 4, emotions are also intentional, something which Ryle overlooks entirely. Sticking to Ryle's limited characterisation of an emotion, we are no clearer

on what role the emotion is playing in motive explanations and why emotions involve the propensities to act in the way that they do.

3.2.2 KENNY ON MOTIVE PATTERNS

Seeking to improve Ryle's proposals, Kenny also aims to address all explanations of action by emotion by distinguishing between behaviour that expresses an emotion and action that is motivated by an emotion, where being motivated by an emotion is having a motive.

In this subsection, I present Kenny's account and emphasise the importance of his contribution of reintroducing the notion of the formal object to our understanding of the intentional nature of emotion. I shall, however, raise an initial problem with his account which is that, like Ryle's account, it cannot accommodate the pure expressive action of Category C.

Similar to Ryle's account, on Kenny's account the role of the emotion in Category A comes out as giving a causal explanation of a manifestation or symptom of an emotion that is an expression of emotion. Without necessarily accepting the details of how emotions manifest themselves, something which must be left to empirical sciences and not to hypothesising about inclinations and inhibitions, Ryle and Kenny are surely correct in identifying the causal role of emotion in Category A. This is so long as we understand 'cause' in a fairly loose sense in order to keep open the possibility that the manifestation just is the emotion rather than being strictly caused by it. In any event, the explanation in Category A is one where the emotion is behind or causes some physical manifestation and does not play a role in the action *qua* rational.

Kenny, however, has an alternative proposal for acting from a motive, which he develops from ideas found in Anscombe (1957). Anscombe distinguishes between three kinds of motives: forward-looking motives, which she also identifies as intentions; backward-looking motives like remorse, pity and revenge; and interpretative motives or 'motives-in-general' like love and fear. While Kenny acknowledges that her distinctions are valuable, he complains that we are still left with the question of 'how do motives explain actions?' (Kenny 1976, 85). Kenny writes:

> The difficulty in giving an account of the concept of motive arises partly from a vagueness in the concept itself. What words count as motive-words? Several answers, equally plausible, suggest themselves. A motive-word, we may say, is a word which is an appropriate completion of the following sentence frames: "He acted out of . . . ", ". . . made

> him do such-and-such", "He did such-and-such because he was…", "His motive in doing such-and-such was…". Unfortunately, the use of these sentence-frames to define motives produces four lists of words which by no means coincide. (Kenny 1976, 85)

Kenny thus seeks to improve both Ryle's and Anscombe's accounts of explanations of action in terms of motives. He contrasts motives with intentions and proposes that acting from either is an exemplification of a pattern where there is a state the agent dislikes, she does something, and ends up in a state she prefers (Kenny 1989, 62). When we give an explanation, we could give it in terms of the initial disliked state of affairs, the preferred state of affairs which is taken to be the upshot of the action, or some combination of the two.

Acting from an intention is one exemplification of the pattern. When acting from an intention, one is acting from a forward-looking reason, and acting in order to achieve something. Reporting an agent's intention is referring to that forward-looking reason. For example, the man flatters the Prime Minister in order to become a bishop. His intention is to become a bishop and he is acting to move from one state of affairs where he is not a bishop to a preferred state of affairs where he is one.

Acting from a motive is another exemplification of this pattern. Unlike a report of an intention which gives the forward-looking reason, a report of a motive 'may either give a backward-looking reason, or exhibit the action as falling under some specific scheme of this general pattern' (Kenny 1989, 63). For example, if I killed the man out of revenge, my motive of revenge refers to my backward-looking reason that the man killed my father. Or, if I roll up my car window because I am afraid of the people milling around at the traffic lights, then my motive of fear exhibits my action as one where I am acting to change the state of affairs where I feel unsafe to something more preferable, feeling safe. I am not necessarily acting with the intention of feeling safe, however.

The agent need not be aware of what motive she is acting from, but she must be aware of the features which make the pattern applicable (Kenny 1989, 63). Othello, for instance, may not be aware that he is acting out of jealousy even if his motive is jealousy. He is, however, aware of himself as acting because his wife Desdemona has apparently betrayed him, a backward-looking reason, or he is aware of his unpleasant feelings and is acting to change them, thus he is aware of a general pattern to his actions.

In this way, Kenny meets the challenge he raises for Anscombe about how motives *explain* action. Motives explain by showing how an action was an exem-

plification of a certain pattern. Motive words which fill in for 'He acted out of…', '…made him do such-and-such', 'He did such-and-such because he was…', 'His motive in doing such-and-such was…' could easily form four different lists, but they would have something in common in that they describe the action as falling under the motive pattern.

We might still worry, however, that Kenny's proposal is not much of an improvement on Ryle's account. The problem that I raised for Ryle was that his account of emotions was unsatisfactory in distinguishing between bodily feelings and emotional feelings and, as a result, I found it explanatorily wanting. In a similar vein, we could ask of Kenny: what is it about emotions which makes them motives? Is it simply how they feel? If so, then his account is no more satisfying than Ryle's. Luckily, Kenny does have more to say about emotions and, in particular, about their intentional nature.

We have already seen that the way in which emotions straddle the mind and the body is a key theme in the history of philosophy and that the Wittgensteinian distinction between the cause and object of an emotion calls for a nuanced treatment of emotions. Yet full appreciation of the way emotions are mental states that are about particular objects which may be distinct from their causes can often seem to be lacking. For example, in *Ethics*, Spinoza gives a limited characterisation of what emotions are about. He writes that '*[l]ove* is nothing else but pleasure *accompanied by the idea of an external cause: Hate* is nothing but *pain accompanied by the idea of an external cause*' (Spinoza 1984, 75). David Hume, in *A Treatise of Human Nature*, also gave a rather limited definition of emotions as sensations of agitations plus belief. For Hume, the belief and not the emotion has an object which it is about. But, as Kenny argues, this entails that the connection between what the emotion is about and the emotion is contingent when it is in fact a logical one. Kenny writes:

> One cannot be afraid of just anything, nor happy about anything whatsoever. If a man says that he is afraid of winning £10,000 in the pools, we want to ask him more: does he believe that money corrupts, or does he expect to lose his friends, or to be annoyed by begging letters, or what? If we can elicit from him only descriptions of the good aspects of the situation, then we cannot understand why he reports his emotion as fear and not as hope. (Kenny 1976, 192)

The logical connection is not between the emotion and its particular object, however, but between the emotion and its formal object. The formal object of a mental state, as noted in Chapter 1, is a medieval notion that Kenny influentially reintroduced into contemporary theorising about emotions. For Kenny, '[t]he formal

object of ϕing is the object under that description which must apply to it if it is to be possible to ϕ it' (Kenny 1976, 189) and '[e]motional attitudes, like other mental attitudes, have formal objects' (191).

An emotion like love is thus not just pleasure accompanied by the idea of an external cause, *à la* Spinoza. It is a mental state whose object is such that there is a description which must apply to it if it is to be the object of love. The external cause, as the emotion's object, must be loveable, affection-inspiring, etc., just like winning £10,000 in the pools must be dangerous, threatening, etc., if it is to be feared.

While the intentional nature of emotions has not been completely overlooked historically—for instance, as Kenny himself notes, Aristotle already defined what would be the formal objects of emotions in *On Rhetoric*, such as when he defines anger 'as a distressed desire for conspicuous vengeance in return for a conspicuous and unjustifiable contempt of one's person or friends' (Aristotle, *On Rhetoric*, 1378a)—crystallising the idea of emotions as having formal objects, as well as the idea that the connection between the formal object and the emotion is a logical one, is an important step in our understanding of emotion as intentional. And because emotions are intentional, an emotion could be a motive because of more than how it feels. If Othello acts out of jealousy, for instance, we know that he does not only have unpleasant sensations, he also sees a threat to Desdemona's affections for him.

While this is an improvement on Ryle's proposal because it introduces the intentional nature of an emotion, Kenny's account faces three problems. The first it shares with Ryle's proposal, which is that it cannot accommodate the pure expressive action of Category C.

When acting purely expressively, such as tearing at the photograph of one's rival or rolling around in the clothes of a lost loved one, the agent is not going from a disliked state of affairs to a preferred state of affairs; nor is she necessarily aware of her action as following this pattern. She does not have a backward-looking reason, although she acts because of something contemporaneous with the action, namely her emotion, and she does not act to change how things are. Yet, she is acting intentionally and the emotion is not simply causing her movements or behaviour in the way it does in Category A. Category C does not fall under either of the explanatory roles Kenny describes for emotion.

The second problem is that the motive pattern, by being too general, misses important differences between ways we can act from an emotion. The third problem is related, which is that by being too general, the motive pattern in fact also misses an important explanatory element of motives. I shall address these two problems together in the next section.

3.2.3 Different ways in which we can act from an emotion

Kenny proposes a pattern of acting-to-change which is supposed to capture what is going on when an agent acts from a motive. An explanation in terms of a motive either reports that there was some backward-looking reason or shows how the action was an exemplification of the motive action pattern. For Kenny, motives include more than those typical motives for which a motive word springs to mind, like revenge or jealousy. He writes:

> 'I'm coming indoors because it's too cold out here' 'They sacked him because he turned up drunk three days running' 'He bought a new suit because the old one was too small'. These sentences obviously have something in common with 'I killed him because he killed my father'. Yet there is no motive-word which occurs naturally in connection with the first three sentences as 'revenge' does with the fourth. (Kenny 1989, 62)

What they have in common, for Kenny, is that they are reports of backward-looking reasons, the action follows the pattern of acting to effect a change from one state of affairs to a more preferable one, and the agent is aware of herself as acting in this way (Kenny 1989, 62). Kenny goes on to say:

> Which backward-looking reasons we will naturally call 'motives' depend on the comparatively trivial circumstances of whether or not we have a name for the specific scheme exemplified. Where we have a common type of undesirability in the pre-action state of affairs, or a common type of desirability in the post-action state of affairs, we assign names to the particular pattern of action, and speak of action out of—for instance—fear, jealousy, or ambition. (Kenny 1989, 64)

If having a name is the only difference between the two kinds of cases, then an account of the nature of motive explanations need not make a fuss over the difference, and a general schema like Kenny's is sufficient. Unfortunately, this is not the only difference. Motive-proper explanations, ones which involve motives like 'revenge', provide fundamentally different explanations to simply describing a pattern according to which an agent acts or reporting that there were backward-looking reasons.

To see this, we must distinguish between at least three ways we can understand action undertaken because of an emotion which fall under Kenny's motive pattern.

In the first way, we are acting *out of* that emotion. For example, Othello murders Desdemona *out of* jealousy. There is a common type of undesirability to the pre-

action state of affairs when someone acts out of jealousy, which is that one sees oneself as losing someone's affections to another. There is also a common type of desirability to the post-action state of affairs, which is that one is no longer losing that affection or is somehow regaining it. Or, take Frank who reports Hawkeye and McIntyre to the captain *out of* anger. He is angry with them and takes steps to get them punished. Feeling offended is a common type of undesirability to the pre-action state of affairs that the motive of anger picks out. These are the examples in Category B3.

In the second way, we are acting *because of* an emotion and are addressing symptoms or behavioural manifestations of that emotion. This is like the examples in categories D1(a) and D1(b) and, in these cases, the emotional experience *qua* an experience is the agent's putative reason. In D1(a), the agent either acts to address the undesirable physical manifestations of her emotion, or acts to induce the desired manifestations. The violinist, for instance, takes a sedative to steady her nervously-shaking hands, but Begbie picks a fight to induce the angry adrenaline rush which he enjoys. In D1(b), the agent addresses the behavioural manifestation of her emotion. Claire the agoraphobe avoids open spaces because they induce undesirable panic attacks in her. Dolokhov dislikes the way his depression manifests in apathy; he actively tries to counteract it by being cruel to others. Acting because of an emotion in this way is not expressing an emotion, neither is the action caused by a corresponding feeling. The agent acts purposively because of the emotion in order to address the state, independent of its intentional nature.

These examples are like the general motive-pattern cases like buying a new suit because the old one no longer fits. Contrast this with the first way, where Othello murders Desdemona *out of* jealousy. If Othello were to act *because* he was jealous in a way like the second, he may rather go to talk to the army counsellor about his unfounded but influential thoughts regarding Desdemona and Cassio. Frank may take anger management classes.

The third way is acting *because of* the emotion as an intentional state, like the examples in Category D2 illustrate. For example, Jackson throws out his former partner's possessions because it gives him relief, but only because his distress is about that particular person and these are her possessions. Alice is indignant, even angry, about the way her boss has treated her colleague; she is edgy and in a state of discomfort about an injustice and acts because of that edginess. If either Jackson or Alice were to act to address the state purely as it is manifested, such as by taking a sedative to remove their disliked physical states, then their actions are successful. This is the second way of acting. But if Alice takes a sedative because she is edgy

about some injustice or Jackson takes a sedative because he is upset about a loss, then their actions are not successful—unless Alice took the sedative so that she could think clearly about what to do in order to fix the injustice or Jackson took the sedative so that he could forget about his fond memories of something lost and sleep. These different success conditions for the action show that the nature of why the agent acts is quite different between the second and third ways, where in the third way that agent acts to address her emotion as an intentional state but not necessarily to address what her emotion is about, except as it would change her emotional state.

But now it seems that someone can act out of or because of an emotion in a variety of ways. Someone can act from the 'motive' of anger because it is her motive-proper, as in B3; or she can act because she likes the way the emotion feels or wants to address the way it affects her behaviour, as in D1(a) and D1(b); but she could also act from anger because of what she is angry about. In all cases, she may have a backward-looking reason or be acting in a way that meets the motive pattern. Yet, the pattern alone cannot be used to explain what the differences are and it is not the case that 'anger' is a name assigned to a common type of undesirability in the pre-action state of affairs or a common type of desirability in the post-action state of affairs.

These differences, which Kenny's proposal fails to capture, indicate a third problem, namely that his proposal overlooks a substantial difference that exists between a motive-proper explanation—acting *out of* an emotion, the first way—and acting *because of* an emotion, the second and third ways.

With the first set of examples of acting *out of* a motive-proper, we learn about the manner in which the agent acted by getting a sense of what was going on in the agent's head, of how she was thinking and feeling when she acted. Reporting that Othello acted out of jealousy, for instance, does not merely identify how his pre-action state of affairs where he sees himself as losing Desdemona's affections to Cassio is undesirable, it also tells us that he was consumed with thoughts of Desdemona and Cassio together, looking for evidence to confirm his suspicions. Frank could have neutrally reported Hawkeye and McIntyre to the captain, but because he acts out of anger we learn that he was exasperated, probably could not concentrate on things other than the offensiveness of the troublemakers, and was desiring their punishment.

Anscombe describes this explanatory element of describing the manner of acting in a motive explanation as giving the 'spirit' of the action (Anscombe 1957, 18). As she notes, saying that I did something out of love cannot be paraphrased to say that

I acted so as to bring about some good for the person I love, such as 'to release him from this awful suffering', because we do not then capture 'the spirit' in which I acted. I could act to bring about some good for the person but not be acting out of love, so much as out of prudence, for instance. The motive of love, Anscombe maintains, does not specify some further goal such as being loving; it just is the spirit in which I act and what makes the act one of love rather than one of prudence. We can bolster what this 'spirit' is by appealing to our understanding of emotions as intentional. The formal object, remember, is what all instances of an emotion type share. So, if the formal object of jealousy is something like resentment of 'a third party for loss or threat to another's affection or favour' (Lazarus 1991, 122), then acting out of the motive of jealousy brings with it the associated spirit where one is alert to evidence of someone's failing affection, preoccupied with the matter, and desires to act because of it. It is because the formal object of anger involves seeing an offence to oneself that one becomes extra sensitive towards the person who has offended you, has typical thought patterns, and so acts in a certain spirit.

In contrast, we do not get the same internal insight with the second set of examples where we act *because of* the emotion. When the violinist takes a sedative because of fear and in order to steady her hands, is her fear exhilarating, even if her hands shake? When Begbie picks a fight in order to have an angry adrenaline-surge, is his anger distressing or attention-consuming even though he craves the adrenaline? These examples do not explain by giving the 'spirit' of the action because we do not get a sense of what was going on in the agent's head, what she desires, or even necessarily how she was feeling, just like we do not get a sense of what was going on in the agent's head when she washed or bought a new suit.

We therefore see that the role of the emotion in Category B3 where the emotion is the agent's motive-proper is quite different to the role of the emotion in Category D where the emotion is the agent's reason because it is what she addresses. Meeting the motive-pattern does not exhaust the explanatory force of identifying a motive-proper such as love because the motive-pattern does not include the 'spirit' of the action and the 'spirit' arises partly because of an emotion's intentional nature. While Kenny rejects Anscombe's motives as too vague, he overlooks this insight of her account.

The intentional nature also plays a role in the other examples in Category B, where the agent forms other beliefs and desires (B1) or attitudes (B2) which in turn explain the action. General motive-pattern explanations, like buying a new suit because the old one no longer fits or the examples in Category D, do not place the same emphasis on the intentional nature of the emotion or its relation to other

mental states. We therefore see that, even while all the examples illustrate action following the same general pattern, the pattern is not sufficient for a characterisation of acting from a motive-proper and runs together quite different ways one might act because of an emotion, and their explanations.

I shall now go through the examples to give more precise explanations in light of this discussion.

In Category B, the emotion plays a role because of its intentional nature. Making use of the concept of the formal object, we can identify what kinds of beliefs, desires and attitudes an agent forms on the basis of an emotion, and explain those beliefs, desires and attitudes using the emotion.

B1. WHERE THE EMOTION IS CLOSELY LINKED TO OTHER BELIEFS AND DESIRES, PERHAPS BY CAUSING THE BELIEFS AND DESIRES

- *Huck Finn failed to turn Jim over to the authorities because of his compassion for Jim.*
  - More precise explanation: *Huck Finn failed to turn Jim over to the authorities because his compassion about someone valuable led him to believe that Jim deserved protection.*
- *Alice challenged her boss out of moral indignation.*
  - More precise explanation: *Alice challenged her boss because her moral indignation about something unjust led her to believe that she ought to challenge the perpetrator of the injustice.*

B2. WHERE THE EMOTION AFFECTS HOW WE SEE A SITUATION, IN AFFECTING WHAT ACTION THE AGENT PERFORMS OR HOW THE AGENT BEHAVES

- *Begbie beats up the man because he is angry.*
  - More precise explanation: *Begbie's anger makes him see innocuous situations as offensive and he beats the man up because he finds him offensive.*
- *Frank kicks up a fuss because he is angry with Hawkeye and McIntyre.*
  - More precise explanation: *Hawkeye and McIntyre often make Frank angry and he is predisposed to interpret their actions as offensive. He kicks up a fuss because he finds them offensive and gets angry with them.*

- *Pierre gives alms to the poor because of his overwhelming love.*
  - More precise explanation: *Pierre gives alms to the poor because his love for Natasha makes him see everything and everyone as good and deserving respect.*

B3. WHERE THE EMOTION IS THE AGENT'S MOTIVE FOR ACTING

- *Frank reports Hawkeye and McIntyre to the commander out of anger.*
  - More precise explanation: *Frank reports Hawkeye and McIntyre in a spirit of anger because of the offenses they've committed towards him.*
- *Othello kills Desdemona out of jealousy.*
  - More precise explanation: *Othello kills Desdemona in a spirit of jealousy because of he thinks that she has committed adultery.*

In Category D, in contrast, the emotion does not play a role in the explanation because of its intentional nature. Perhaps, like in D2, the intentional nature is important in determining what kinds of actions can be taken which will effectively address the state that the agent is in. However, such a role is not in terms of being connected to certain patterns of thought and attitudes. We can therefore give the following more precise explanations:

D1(A). ADDRESSING THE SYMPTOMS OF THE FEELING OR MANIFESTATION OF THE EMOTION

- *The violinist took a sedative because she was nervous.*
  - More precise explanation: *The violinist took a sedative because her hands were shaking nervously.*
- *Begbie picks fights because he enjoys being angry.*
  - More precise explanation: *Begbie picks fights in order to simulate situations where he gets an adrenaline rush which he enjoys.*

D1(B). ADDRESSING THE WAY THE EMOTION IS MANIFESTED IN BEHAVIOUR

- *Claire avoids open spaces out of fear.*
  - More precise explanation: *Claire avoids situations where she feels overwhelming fear because the fear is unpleasant.*
- *Dolokhov challenges Rostov to gamble away money he does not have out of cruelty.*
  - More precise explanation: *Dolokhov challenges Rostov to gamble away money he does not have because he seeks to break his unwanted feeling of apathy by being cruel.*

D2. NOT JUST ADDRESSING THE FEELING OR MANIFESTATION

- *Jackson throws out his ex-partner's belongings out of grief.*
  - More precise explanation: *Jackson throws out his ex-partner's belongings because he is upset about his loss, and throwing out the belongings addresses his sadness by removing memories of what he is sad about.*
- *Alice challenges her boss because she is indignant.*
  - More precise explanation: *Alice challenges her boss immediately because the edginess of her moral indignation about the injustice spurs her to act in challenging the injustice.*

3.3 CONCLUSION

Kenny allows that emotions are intentional and makes use of the notion of the formal object, so why does he seem to overlook the full import of his positive proposal of emotions as intentional? By tentatively proposing an answer, I shall close this chapter by laying out the conception of emotions as providing access to reasons which I would like to explore in the rest of this thesis.

Kenny ultimately says very little about the formal object of an emotion and his focus is primarily on how we develop our concepts of specific emotions. He argues that we develop our concept of a specific emotion and learn to identify the feeling of that emotion from the way in which its object and context, the symptoms or manifestations, and action tend to co-occur (Kenny 1976, 99; Kenny 1989, 64-65). He is not concerned with the questions of *why* the contexts which co-occur

with certain manifestations include patterns of thought, and *why* they tend to lead to certain kinds of actions, over-and-above the fact that they do co-occur. His focus on the formal object is thus on its role in the development of our emotion concepts and not so much on what emotions are about. Tentatively, it is because of this difference in focus that Kenny does not explore the full explanatory import of his positive proposal.

However, the link between the emotion and the action is not just one of co-occurrence. There is a reason why certain emotions lead to certain actions, and that reason lies in the intentional nature of emotion. If we approach motive explanations from this direction, then we do find a difference between ways in which behaviour and action fit Kenny's patterns.

If this is right, then we must look closer at the intentional nature of emotions. Doing so opens up the possibility for acting purely because of what an emotion is about or represents, and not because of the 'spirit' or other beliefs or attitudes we might form, as in Category B, or by being the agent's reason because it is a state to address, as in Category D.

Consider some of the examples in Category E. Richard buys flowers for Katherine out of love. He is not acting for a backward-looking reason and he is not caused to buy flowers because of the way his love feels. He chooses to buy them, and does so because he loves Katherine. Natasha, distraught at the death of Prince Andrey, pulls her life back together because she loves her mother. She is acting to bring about some change, but that is not what acting out of love is in this context. In both of these cases, there is not a clear sense that the agent is acting to change things or acting because of some backward-looking reason. Rather, both Richard and Natasha act to express their emotions, but in a way different to Category C's purely expressive actions, which I shall examine in Chapter 5. They act because they value someone, and this plausibly is the reason why they act. Is it a reason? If yes, what role does the emotion play in the action? Does Richard, say, still need to form a belief that Katherine is someone he values if he is to act for a reason, or can he act directly on the basis of an emotional experience? Is doing so reasonable or rational? These are questions I shall answer in Part III.

To summarise what has been covered in this chapter, I began by identifying three key themes found in historic thought about emotion. These were: (i) the emotions, as passions, were seen to be passive and opposed to activity, like perceptions; (ii) the emotions were seen as states that straddle both the mind and body; and (iii) the emotions were given a motivating role in action. I proposed that these themes are not at odds with ideas central to the contemporary pro-emotion consensus which

supports the compatibility of emotion with reason. I also introduced two distinctions that have been made in twentieth-century philosophy, a distinction between the cause of and reason for an action and a distinction between be cause and object of an emotion. I argued that these distinctions allow us to explore a more positive role for emotion in action, and that the three historic themes actually support our doing so.

I then looked at two proposals for the role of emotion in action, Ryle's and Kenny's, and argued that neither proposal successfully applied to all the categories of explanation which I identified in Chapter 2. This gives us a *prima facie* reason to look further at what role emotions can play. I further argued that Kenny's positive proposal is too general and the kind of distinctions it misses are ones which rely on the fact that emotions are intentional states which lead us to form certain beliefs and desires and adopt certain attitudes. If this is the case, then the conceptual space for emotions to be involved in action purely in light of their intentional nature is opened up, including the possibility that acting on the basis of an emotional experience could be acting in light of a reason and that an emotion explanation could rationalise action in some way.

Before being in a position to spell out fully what those roles are, however, we must come to grips with what emotions must be if they are intentional, as well as be clear on what acting for a reason and rationalising explanations involve. Thus, in Part II, I set out in Chapter 4 by examining some of the most influential contemporary accounts of emotion and motivate adopting a minimally cognitive account such as one making use of an analogy with perception and, in Chapter 5, I examine acting for a reason and rationalising explanations.

PART II

CHAPTER 4

ACCOUNTS OF EMOTION

In Chapter 1, I drew attention to the diversity of things which we include under our concept of emotion, including agitations, occurrent emotion states, emotional attitudes and moods. I identified two features shared by this range to greater and lesser extents, but which are central to our emotion concept. These features were: (i) that emotions have an intentional nature; and (ii) that emotions have an affective nature including an intentional quality to the way they feel. In Chapter 2, I introduced a range of examples of the different roles emotions can play in explanations of action and, in Chapter 3, I argued that there is conceptual space for an emotion to play a role in action and action explanations beyond being a motivating state and in virtue of its intentional nature. In the next two parts of my thesis, I explore how emotions can play such a role and develop the view that emotions can have a reason-providing role in action and explanations of action. In Part II, in chapters 4 and 5, I give necessary background on emotions and action explanations. In Part III, in chapters 6 and 7, I show how emotions can provide access to reasons and how we can act in light of those reasons when acting on the basis of an emotional experience.

Because so much turns on the intentional nature of emotion, more needs to be said about what emotions are such that they have this nature. This is the task of the current chapter. After the discussion of Chapter 1, it may seem that providing a comprehensive analysis of what the necessary and sufficient conditions are for something to be an emotion is a daunting task—if not an impossible one. So, in this chapter, I focus on occurrent emotion states. Occurrent emotion states are both clearly affective and intentional; they stand out as the paradigm kind of emotion. They have also been the primary focus of accounts of emotions and so it makes sense to limit my attention in a similar way.

There are other reasons why I focus my attention on occurrent emotions. Because my ultimate focus is on action and explanations of action, and on whether

emotions can provide access to reasons for action, I set aside agitations. Agitations, by the definition given in Chapter 1, are just reactions. If they result in some kind of behaviour that is not merely a compulsive reaction, they are by definition no longer agitations but are occurrent emotions. If the behaviour is part of the reaction, then the role of the agitation in explaining that behaviour is simply causal. As pure agitations, they are not relevant to my interests.

I set aside moods because they do not have stable particular objects. Occurrent states and attitudes are directed at particular objects and, like agitations, are responses to those objects. If we act from an emotion, we act in response to something that we think obtains and we would explain the action in terms of that response. A mood, in contrast, is not itself always a response to *particular* things in the environment. Because of this, acting from a mood is *prima facie* quite different to acting from an emotion and deserves its own treatment which is beyond the scope of this thesis.

Finally, I largely set aside attitudes, for three reasons. First, attitudes are much like character traits and many character traits will involve attitudes, such as having the character trait of being a loving person is having a loving attitude towards others.[1] Attitudes and character traits both involve more than a single experience, instance or state and, like moods, deserve their own treatment. Second, because attitudes involve more than a single instance, their role in action explanations is plausibly most like the roles illustrated in Category B where the emotion provides an explanation of other mental states the agent has. For the rest of this thesis, however, I wish to explore the explanations in Category E. Third, attitudes are still similar to occurrent emotions in many respects, such as by having particular objects and, as we saw in Chapter 1, they share a similar intentional nature and partly involve dispositions to experience certain occurrent emotions. Attitudes may thus be illuminated if we have a good account of occurrent emotions. So, while attitudes are not my main focus, addressing occurrent emotions will hopefully provide insight for attitudes as well.

Even limiting my focus to occurrent emotions, I will not address many of the accounts of emotions that have been proposed in recent years.[2] I have chosen to focus on the accounts and issues that I have in order to illustrate the intertwinement of the affective and intentional elements in an occurrent emotion, and to indicate

[1] One key difference between an attitude of love for X and the character trait of being loving is that an attitude of love has a set particular object, X, while the character trait need not.

[2] Useful and comprehensive overviews have been given by others already. I have found Deonna and Teroni (2012) particularly helpful, as well as the more compact sections in Deigh (2010) and Brady (2013)

what we must be minimally committed to if occurrent emotions are intentional.

In what follows, I therefore focus on influential accounts of the twentieth century and break them down into three general classes of views. These classes need not be mutually exclusive; I choose to present them this way in order to make their particular emphases clear.

The first class is of views that highlight the feeling side and non-cognitive elements of an emotion, such as its bodily manifestations and sensations. I side with those who argue that such accounts are unable to do justice to the intentional element of emotions and that some cognitive involvement is required in an account of emotion. The second class is of views that highlight the cognitive elements of an emotion, identifying them as beliefs or judgements. I side with those who argue that such accounts are either overly cognitive or are trivial. The third class is of views that propose a weak kind of cognitive involvement, such as the currently-popular perceptual model of emotion. Such views draw on perception as an intentional state that is not belief- or judgement-like as a model for emotion. Even while I remain neutral about the extent to which the perceptual analogy can be pushed, I shall argue that the weak cognitive approach of the third class is the right kind of approach to take and the perceptual model a good illustration.

With this in mind, the underlying argument of this chapter can be summed up in the following way: emotions are about things other than our bodies and in order to be that way, some kind of minimal cognitive involvement is necessary. Two corollaries of this conclusion are important for the rest of my thesis. First, if emotions are about things other than our bodies, then they give us *information* of something other than our bodies, information that could make up facts in light of which we act. Second, if some kind of weak cognitive involvement is necessary and is not a belief or a judgement, then since emotions provide us with information of something other than our bodies, they can play a role in action which is different from that of a belief or judgement.

4.1 VIEWS WHICH FOREGROUND NON-COGNITIVE ELEMENTS

One of the most common ways to think about emotions is to conceive of them as feelings. We ask people how they feel and expect reports of their emotional states. You feel angry, sad, surprised, jealous; you also feel flushes of anger, pangs of love; your heart aches or your stomach turns. Given the importance of emotional feeling and the way emotions manifest in the body, it is no surprise that some accounts of emotion take seriously the idea either that emotion just is a class of feeling, or

that emotion just is the bodily change. According to such accounts, emotions are not constituted by beliefs, judgements or other cognitive activity, nor are they necessarily caused by beliefs, judgements or other cognitive activity. These accounts foreground the non-cognitive elements of an emotion—some even discount or eliminate the cognitive element—and can accordingly be called non-cognitive accounts of emotions.

In this section, I begin by giving an overview of two specific versions of non-cognitivism. The first is an instance of what has been called the 'feeling theory' of emotion. According to feeling theories, an emotion is the feeling of bodily change. In the late nineteenth century, the psychologists William James and Carl Lange independently developed similar theories and we now talk of the 'James-Lange' theory of emotion as the ancestor of modern-day feeling theories. I shall focus on James' version as it is more dominant in the literature and has undergone a revival in the last twenty-odd years. According to his version, emotions just are feelings of bodily changes, identified as changes in the viscera. After introducing non-cognitive theories via James' theory, I focus on a modern-day version of a feeling theory although, strictly speaking, this is not so much a 'feeling' theory as a more general non-cognitive theory. This is the proposal that emotions are the bodily changes themselves—including neurological changes as well as visceral changes—which goes together with the suggestion that the feeling of an emotion is the feeling of bodily change. This is a theory put forward by the neuroscientist Antonio Damasio.

Following others, I argue that pure non-cognitive theories like these are unable to give due justice to the intentional element of an emotion. This is because emotions are about more than the body and, therefore, they must involve some kind of cognitive element. I next consider Jesse Prinz's modified feeling theory, inspired by James' theory, as a response to the intentional challenge, but argue that it fails to meet the challenge. I conclude that an adequate account of emotion must be able to capture the intentional element, and that keeping an account of emotion purely non-cognitive fails to do so.

4.1.1 BODILY CHANGES AND THEIR FEELINGS

> My theory ... *is that the bodily changes follow directly the* PERCEPTION *of the exciting fact, and that our feeling of the same changes as they occur* IS *the emotion.* (James 1884, 189-90)

This is how James describes his own theory. According to him, the emotion occurs directly on perceiving some exciting fact and is not mediated or caused by a belief or judgement, etc. Neither is the emotion constituted by any beliefs or judgements, etc. It just is the feeling of bodily changes, which James identifies as the feeling of visceral and muscular changes. His account is non-cognitive.

To see how the account works, consider an example. I see a car out of control and heading towards me at great speed. This is the exciting fact. My perception leads to my heart rate increasing, my palms sweating and my muscles tensing up. These are visceral and muscular changes. And the feeling of my increased heart rate, sweaty palms and tensed muscles is my emotion, fear. Because my emotion is the feeling of the bodily changes, I would no longer experience the emotion if there were no longer bodily changes to feel.[3]

In 1927, Walter Cannon wrote an attack on the physiological evidence given by both James and Lange. With regards to James, Cannon argues that the physiological changes in the viscera and innervation of muscles that James identifies are not the sources of the quality of feeling of emotions. He cites evidence that the same visceral changes occur both in different emotional states and in non-emotional states. Nevertheless, we are able to distinguish between emotions and non-emotions and between different emotional states on the basis of the way they feel. Cannon thus argues that the bodily changes are not enough to distinguish between different emotions, and concludes that 'the responses in the viscera seem too uniform to offer a satisfactory means of distinguishing emotions which are very different in subjective quality' (Cannon 1927, 109). A further conclusion that he draws is that, if James' theory were true, then we would expect 'not only that fear and rage would feel alike', because they have similar visceral changes, 'but that chilliness, hypoglycaemia, asphyxia, and fever should feel like them' since they too have similar visceral changes (100). This, clearly, is not the case. We can and do distinguish between emotions and non-emotions and between different emotions on the basis of the quality of how they feel.[4]

[3]The focus of James' account is on occurrent emotion states and so it would not be an objection that there are such things as emotional attitudes which do not have bodily profiles.

[4]This is not to say that we are always able to do so or that we are always right in what emotion

By undermining the physiological support for the theory, Cannon does not strike a decisive blow against the core idea of the feeling theory, which is that an emotion is constituted by a bodily change of some sort and in some combination. More nuanced versions of feeling theories have developed the kind of bodily change involved in order to address Cannon's kind of objection. The neuroscientist Damasio (1994), for example, proposes a modern version of a feeling theory. On his theory, the relevant changes are not limited to bodily changes; they include neurological ones. If this is the case, then the difference in emotional feeling may arise from differences in neurological change, even if the bodily changes are the same.

The most important difference between James' and Damasio's feeling theories, however, is that Damasio identifies the emotion not with the feeling of bodily changes but with the bodily changes themselves. For Damasio, emotions are complex, automatic and unconscious bodily reactions to certain stimuli. When we are afraid, for instance, our hearts begin to race, our muscles contract and there are changes in our brains. This is the emotion and we need not be aware of it. Only when we become aware of the physical changes do we experience the feeling of fear, even though we are already in a state of fear. The emotion is the bodily change and the feeling of the emotion is the feeling of the bodily change.

There is something quite counterintuitive to non-cognitive theories like these. We usually think that our feeling an emotion is an immediate response to the stimulus. But for both kinds of non-cognitive theory, the feelings which we ordinarily classify as the emotion are instead direct responses to the bodily changes. On James' version, rather than feeling fear in response to the oncoming car, I feel fear in response to my bodily changes, brought on by the car. As James notes, this is in reverse order to our commonsense understanding of emotion, which would have me feel fear in response to the car and then undergo bodily changes. Damasio's view is no less counterintuitive. For him, even while the emotion as bodily change is a direct response to the stimulus, my conscious feeling of fear is a response to the bodily changes and not a response to the car.

The counterintuitive appearance may be just that, an appearance. The conscious emotion (James) or emotional feeling (Damasio) is still a response to the stimulus, albeit mediated by bodily change. Chronologically, the emotion as feeling of bodily change (James) or the emotional feeling of those changes which are the emotion (Damasio) can be concurrent with the changes as they happen. The emotion (James) or emotional feeling (Damasio) is just not a conceptually immediate response.

---

we identify.

There is a deeper problem, however, which is what Cannon's objection brings out. It is that bodily feeling or feeling of more general neurophysiological change, alone, cannot account for the particular *quality* of emotional feeling.

As I discussed in Chapter 1, we distinguish between emotions and other bodily feelings as well as between different kinds of emotions because of the way they feel. The way they feel could be distinguished either broadly as 'positive' or 'negative' or more nuanced, such as between two negatives like anger and fear. I also discussed how the feeling of an emotion is the feeling of a state that is directed or targeted at objects other than one's body. If an emotion or emotional feeling is fundamentally only a feeling of bodily change, the feeling theorist has the explanatory burden of showing how emotions have a qualitative nature which sets them apart from other kinds of feelings, explaining how we make the distinctions that we do make, and explaining the way emotions feel like they are directed at things other than one's body. In the next two subsections, I shall explore this burden and see if it can be met.

4.1.2 Accommodating the quality of feeling

We do occasionally confuse emotions with non-emotions and different emotions with each other, something which does speak in favour of there being similar bodily changes. Despite this, we usually do not confuse them and, more importantly, we are able to identify and correct confusion. This suggests that there is a crucial difference, a difference which cannot lie in different bodily profiles since different emotions and emotions and non-emotions seem to have the same profiles.

In support of this idea, Stanley Schachter and Jerome Singer (1962) studied physiological arousals more broadly construed than just visceral changes. In their study, Schachter and Singer manipulated the physiological arousal of a group of subjects by giving them an adrenaline injection. By doing so, they induced the same bodily state in the subjects. Schachter and Singer then introduced into the group a 'euphoric' or an 'angry' stooge who acted euphorically or angrily, respectively, to test whether they could manipulate the emotional self-reports and behaviour of the test subjects. The subjects in a group where the stooge acted euphorically tended to report more positive emotions than the subjects in a group where the stooge acted angrily. From these results, Schachter and Singer conclude that '[g]iven a state of physiological arousal for which an individual has no immediate explanation', an arousal induced by the adrenaline injection, 'he will label this state and describe his feelings in terms of the cognitions available to him' (Schachter and Singer 1962,

398), where the cognitions were influenced by the stooge's actions; they also conclude that the same feelings of physiological arousal could be labelled as different emotions depending on the additional cognitions.

This study suggests that physiological changes are not enough to determine what individual emotions we feel and experience. Even if some arousal is important and even if there are subtle differences in the bodily description for different emotions which we have not yet identified, Schachter and Singer's study suggests that the emotion we experience is in part determined by something beyond the feeling of the body state, something like a cognition or an interpretation of that feeling within some context, and not just by the bodily changes and the feeling of the changes, as James and Damasio suggest.

The non-cognitivist could respond that the results only show that cognitions sometimes cause emotions, not that they always do, nor that they are constitutive of the emotion itself. After all, the subjects start in the same state of arousal but the self-reporting euphoric and angry subjects may have different bodily profiles corresponding to their self-reported emotional states by the end. Neither does being caused by a cognition in one instance entail that emotions are always caused by cognitions. In addition, someone could experience her bodily feelings, wrongly classify them, and yet still be experiencing a different emotion. Interpreting the feeling one way does not make the experience whichever emotion the report says it is.

These responses, though they have some force, fail to meet the point that the study demonstrates, namely that the bodily arousal is not enough to generate distinct emotions. It is the bodily arousal plus an interpretation of the feeling within a context which leads the subject consciously to experience an emotion. Whether her interpretation is correct or not, some interpretation is required and the emotion is not a response solely to bodily changes. Later studies provide further data to support the claim that bodily changes alone are insufficient for an emotion, even when manipulating processes in the brain to constitute one emotion rather than another.[5]

Where does this leave a refined feeling theory like Damasio's, where an emotion just is the somatic change? Damasio's theory is compatible with Schachter and Singer's findings because, if there is no physiological arousal in the first place, then there is no arousal to be interpreted and no corresponding emotion to be reported. The physiological arousal would be Damasio's emotion and what is reported would

[5]See Barrett et al. (2007) for a summary of a range of studies and the argument that stimulating regions of the brain that are supposedly the base of certain emotions rarely produces experiences of discrete emotions.

be the feeling of the emotion. However, the particular feeling of an emotion is the current focus, and so the relevant part of Damasio's theory to be examining is the emotional feeling. And here we see that, even if we were to accept Damasio's definition of an emotion as bodily changes and even if we were to accept that each emotion has a distinct physiological or neural profile, the feeling of the emotional experience still depends on factors other than that profile. What Schachter and Singer show with their manipulation of conditions is that the feeling and qualitative nature of an emotion is in part dependent on cognitive factors. If we are seeking to explain the conscious experiential side of an emotion, the qualitative nature of emotional feelings, something other than the bodily change has to come in.

As discussed in chapters 1 and 3, one of the major ways in which bodily sensations are distinct from emotional feeling is that emotional feeling involves the state being directed at objects external to the body whereas mere bodily feelings do not; I called this the 'intentional quality'. The challenge for the non-cognitivist is to show how this can work. If the non-cognitivist is to meet the challenge illustrated by Schachter and Singer's experiment, then she needs to show how emotions have their intentional quality, a quality which distinguishes both between different emotions and between emotions and mere bodily sensations, and do so in a way which does not introduce elements that detract from her core claim, that emotions are non-cognitive. Prinz (2004) takes up this challenge and argues that the bodily changes are the means by which one tracks and represents external stimuli.

4.1.3 PRINZ AND EMBODIED APPRAISALS

Prinz (2004) develops a neo-Jamesian feeling theory, what he calls an 'embodied appraisal' account of emotions. It is neo-Jamesian because he argues that emotions are fundamentally feelings of bodily changes; in particular, he argues that emotions are perceptions of bodily changes. It is an 'embodied appraisal' account because he argues that emotions represent an evaluative property through changes in the body, but are not thereby cognitive. This evaluative property that emotions represent is the formal object, what Prinz defines as 'the property in virtue of which an event elicits an emotion' (Prinz 2004, 62). He proposes that Richard Lazarus' 'core relational themes', some of which I introduced in a table in footnote 13 in Chapter 1 when discussing the formal object, are good candidates or approximations for what emotions represent, namely the formal object.[6] For example, sadness represents

[6]He rejects other details of Lazarus' account of emotion, however, such as Lazarus' key idea that the core relational themes correspond to inner judgements, and are not simply external conditions as Prinz will maintain (Prinz 2004, 65).

and is elicited by the property of being a loss; fear represents and is elicited by the property of being dangerous; anger represents and is elicited by the property of being a demeaning offence; love represents and is elicited by the property of inspiring affection. While Prinz uses both 'core relational theme' and 'formal object' to refer to what an emotion represents and is elicted by, I shall use 'formal object' in order to be consistent with my own terminology.

In order to flesh out how Prinz's account works, I start by introducing Prinz's idiosyncratic definitions of cognitive acts, cognitions and perception, before turning to his understanding of representation and of dual content. Because my aim in this section is to see if a plausible account can be given of the intentional quality of emotions without detracting from the non-cognitivist's core claim that emotions are non-cognitive, I accept these background commitments. I then turn to show how Prinz's account fails by its own lights and end by generalising my conclusion regarding Prinz to all approaches which deny that emotions are cognitive in any important respect.

Prinz has idiosyncratic definitions of cognitive acts, cognitions and perception. 'Cognitive states and processes', Prinz writes, 'are those that exploit representations that are under the control of the organism rather than under the control of the environment', meaning that, in cognitive states and processes, the representation is activated by or maintained in working memory (Prinz 2004, 45-46).[7] A cognitive act is one which results in a cognitive state and makes use of a cognitive process. The representations that are under the control of the organism, at least in a dispositional sense, are cognitions. Cognitions include representations that are concepts and the control is dispositional because it is possible for a cognition to be produced automatically and not be intentionally brought forward from memory, which are cases of what Prinz calls 'unthought thoughts' (46). This might happen, for example, on seeing a dog and forming a thought of a dog. In such a case, we do not draw on the cognition actively or intentionally but we do have dispositional control of it because we can go on to think about dogs or identify what we are seeing as a dog.

In contrast to a cognitive state or process, perception is a dedicated input mechanism. Perceptual representations arise from 'out of the brain' (Prinz 2004, 222), meaning that they result from external inputs and are not under our control. Nev-

[7]This is idiosyncratic because cognitive states are often not thought of as under our control. We cannot believe simply on the basis of choosing to, for example. By 'under control', Prinz really means that we are drawing on concepts from memory. Thanks to Sasha Vereker for raising this idiosyncracy and helping me to clarify Prinz's view.

ertheless, perception can also produce cognitions, as we see in the case of the 'unthought thought' of a dog. Perceptual representations and cognitions, however, are different things. Imagine that you see a dog and automatically form a 'dog' thought such as, 'that is a dog'. The content of your thought is made up of a cognition but it is not the result of a cognitive act. You could also just see something which looks like a dog and form no thought about it. You still have a representation of what you are seeing, but this time it is a perceptual representation that is not a cognition.

With these background definitions to hand, we can turn to Prinz's embodied appraisal account. To flesh out how it works, I shall have to say more about Prinz's understanding of representation and of concepts and their dual content.

Core to Prinz's account is the idea that emotions track and represent their formal object by tracking bodily changes. Prinz makes use of Fred Dretske's (1981, 1986) causal notion of representation, according to which something is a representation of an X if it is reliably caused by and co-varies with Xs. Emotions, according to Prinz, are reliably caused by and co-vary with their formal objects. The formal objects, however, are the indirect causes of the emotions. The direct causes which the emotion also co-varies with are changes in the body. Nevertheless, the emotion functions to represent the formal object and not the bodily change because it co-varies with the formal objects of which the bodily changes are simply the appearances.

In order to support this, Prinz draws on his theory of dual content for concepts, developed in Prinz (2000) and Prinz (2002). There, he aims to develop a strongly empirical theory of concepts, inspired by John Locke's ideas of nominal and real essences. According to his theory, concepts have a dual content that is made up of representations of appearances and representations of essences, where the essence is the property in virtue of which something is the thing it is and has the appearances it has. The representations of appearances make up the nominal content of a concept, and the representations of essences make up the real content.

For example, a concept of a dog has both nominal and real content. It represents the doggy features which make up the appearance. It also represents the doggy essence—its doghood—which makes a dog a real dog and have the appearance that it does, perhaps something like a genome. Despite this dual content, our concept of 'dog' functions to represent real dogs and not just doggy appearances because real dogs have their doggy appearances in virtue of having the doggy essence. In the case of emotion, the essence of an emotion is its formal object, the evaluative relations, and the appearance is how that essence is manifested in our bodies. For Prinz, the bodily changes that occur are ones which have an evolutionary history for preparing

us for responses of the right kind (Prinz 2004, 68-69). For example, physiological priming for flight when there is danger would be evolutionarily advantageous. So, in virtue of some evaluative relation obtaining, certain bodily changes are what they are, and emotions function to represent the formal objects which are the real essence.

We thus see how Prinz argues that, as perceptions of bodily changes, emotions are also perceptions of the formal objects much like perceptions of the appearance of a dog can also be a perception of a real dog. Because perception is not a cognitive act on Prinz's characterisation of cognitive acts, neither is an emotion. Even those who endorse cognitive accounts of emotions can agree with this, however, because they could accept that an emotion does not actively involve bringing forward a representation from memory, but maintain that an emotion still involves concepts. Prinz therefore needs an even stronger claim, which he does in fact make: emotions do not necessarily involve *cognitions* (Prinz 2004, 49-50). It is this last claim which Prinz needs to show is true but, as I shall now argue, he fails to do.

If emotions are perceptions, then the representations involved in an emotion must not be under the organism's control; that is, they must not be brought forward from memory. This is according to Prinz's own definition of perception. But they also must not involve cognitions. So, if emotions are perceptions not only of bodily changes but also of the evaluative relations that make up the formal object, then they represent the evaluative relations without cognitions. However, emotions cannot represent the formal object without cognitions.[8] This is because we need cognitions in order to avoid the disjunction problem facing causal accounts of representation like Dretske's, on which Prinz bases his own account of representation.

The general problem is this.[9] If a state represents what it is reliably caused by and co-varies with, then it seems that there cannot be cases of misrepresentation. This is because a state representing X may be activated when either an X or an *X-lookalike* is in the environment and is therefore reliably caused by and co-varies with both Xs and *X-lookalikes*. Such a state does not represent X but rather the disjunction of *X or X-lookalike*. So, the state cannot misrepresent X when there is an *X-lookalike* because it actually represents both Xs and *X-lookalikes*. However, we want a theory of representation to allow that there can be misrepresentation and

[8]Mikko Salmela (2011) provides arguments for a comprehensive rejection of Prinz's position. My objection is similar in kind to Salmela's second objection, which is that emotions still involve cognitions.

[9]See Jerry Fodor's 'Semantics: Wisconsin style' (1984) for the coining of the term 'the disjunction problem', as well as Dretske (1981, 1986), Fodor (1989) and Millikan (1989) for the problem and a selection of proposed solutions.

to be robust enough to rule out *X-lookalikes*. Causal theories are thus found wanting.

Consider seeing a dog. We want to say that my perceptual state represents a dog. However, my perceptual representation of a dog is also activated when I see a wolf on a dark night. My perceptual state of a dog tracks both dogs and wolves-on-dark-nights, and therefore represents the disjunction of *dog or wolf-on-a-dark-night*. If so, then when I think that I see a dog but instead see a wolf-on-a-dark-night, I am not misrepresenting what is there.

The disjunction problem is a problem for Prinz not just because it is a general problem for his causal account of representation, but particularly for his account of emotions because of his dual-theory of content. A perceptual representation of the appearances of a dog on a dark night and a perceptual representation of the appearances of a wolf on a dark night are identical, yet only one of them will track the real essence of doghood and represent a real dog. The perceptual representation, as is, represents a disjunction of something that is either a real dog or a dog look-alike but it does not represent the essence of doghood because the disjunction *dog or wolf-on-a-dark-night* does not have an essence that is the same as doghood. Similarly, an emotional experience could represent either bodily changes that are the appearance of evaluative relations, or it could represent bodily changes that are not. But representing the disjunction is not representing the essence, the formal object.

There have been various proposals for solving the disjunction problem, such as Jerry Fodor's theory of asymmetrical dependency (Fodor 1989), and Prinz puts forward his own in Prinz (2000). There, he appeals to an observational and learning process whereby we form concepts, what Prinz calls 'proxytypes'. In forming concepts, we fine-tune our representations until they only track Xs and are only activated by an X, thereby only representing X and not a disjunction. In this way, the representation of 'dog' comes only to represent dogs and not the disjunction of *dog or wolf-on-a-dark-night*.

But note that Prinz's solution to the disjunction problem involves the development and use of concepts, which are forms of cognitions. The concept and not the perceptual representation is what represents both the appearance and essence. If so, then a perceptual representation does not represent the essence. Simply seeing doggy features does not represent doghood, and simply feeling bodily changes does not represent evaluative relations.

Perhaps Prinz could respond that, once we have developed the proxytype or concept, our capacity for the relevant perceptual representations too will be sufficiently developed. This response is not open to him, however, because it undermines the distinction he wants to make between perceptual representations and cognitions.

Above, I discussed Prinz's example of seeing a dog. If you identify the perceptual representation as being of a dog or have an 'unthought thought' of a dog, then you have a cognition. If the perceptual representation is not to be a cognition and is to remain solely the result of external stimuli, then we cannot draw on a concept from memory to identify that it is a dog and not a disjunction of *dog or wolf-on-a-dark-night*. Again, representations of doghood are not perceptual representations alone.

A similar story must be given for emotions. In order for an experience to represent the evaluative relations, it must track the bodily changes that are the appearances of the evaluative relations and not the bodily changes that are not. But in order to do so, we need to learn and observe when the correct kind of appearances obtain. That is, we need to develop the relevant kind of concept or cognition. If our experience represents the essence, the formal object, and not just the appearance, our experience is making use of a cognition, which is exactly what Prinz wants to deny. He therefore fails to show how emotions can represent the formal objects without cognitions.

My discussion of Prinz has been solely in his terms, which are controversial in their own right. So, my rejection of his account does not sound a death-toll for all non-cognitive theories. Nevertheless, there is a general point that can be drawn out, which is that the formal object is, quite simply, not the same as bodily changes. Experiencing the bodily changes that are associated with an emotion rather than bodily changes, *simpliciter*, must be mediated by some cognitive state. That cognitive state may just be one that is triggered in the right environment and so it does not follow that emotions will be the same as beliefs or judgements. It only follows that emotions are not purely non-cognitive. I therefore now turn to examining cognitive approaches to emotions.

4.2 VIEWS WHICH FOREGROUND COGNITIONS

Emotions are about things other than one's body, giving them their particular intentional quality. It is this external focus of emotions which the non-cognitive accounts I have discussed have struggled to capture. In this section, I go from one extreme to the other, and introduce cognitive accounts which propose that emotions are constituted by either beliefs or judgements. While such accounts are able to accommodate the intentional element of emotions, I argue that they do so in a way which is either too cognitive or trivial.

In section 4.2.1, I begin by looking at a representative account of an approach

which analyses emotions in terms of beliefs and desires in order to explain how emotions have the intentional nature that they do. I shall argue that such an approach is problematic because it over-intellectualises emotions. In section 4.2.2, I shall look at other versions of cognitive accounts, judgement accounts. While such accounts are more nuanced and do not require that emotions involve belief, they do draw on the idea that emotions involve assenting to the way things appear. Nevertheless, such accounts also run into problems, most crucially that they are ultimately trivial. In section 4.3, I therefore turn to a third alternative, which is that emotions are weakly cognitive in a way much like perception is.

4.2.1 BELIEF-DESIRE ACCOUNTS

Early contemporary attempts to capture the intentional aspect of emotion in philosophy did so in terms of beliefs and desire. As an example of this kind of approach, let us look at Joel Marks' (1982) account of emotion.

Marks argues that emotions are belief-desire sets characterised by a strong desire.[10] Joan's fearing the dog, for example, is the same as Joan's believing that *the dog is threatening to bite me* and strongly desiring to avoid it. The desire is supposed to capture the affective and motivating element of an emotion; the belief the intentional element that the dog is threatening to bite her; the two together the evaluative element: the dog is threatening to bite and Joan desires to avoid it. In terms already familiar to us, the belief has the content that *the dog is threatening to bite me* and the *particular object* of the emotion is also the object of the proposition which makes up the content of the belief; in this case, it is the dog. The *formal object* is the full evaluative content of what is ascribed to the particular object. In this example, it is that *the dog is threatening to bite me and I am desiring to avoid it*.

Marks requires that the desire is a strong desire in order to distinguish between belief-desire pairs which are emotions and belief-desire pairs which are not. The difference between the emotion and the non-emotion, for Marks, is not in content but in strength of desire. The belief-desire set in turn causes certain physiological changes which characterise the bodily feeling of emotion.

Such accounts of emotions have been roundly rejected. The most compelling objection is that such an account over-intellectualises emotions. The content of an emotion, on this kind of view, is the content of a belief, which is a proposition. By being afraid, the content of Joan's emotion is that *the dog is threatening to bite*

[10]For other belief-desire combination accounts, see John Searle (1983) and O. H. Green (1992).

me, a proposition which is in turn about the dog. This means that emotions require grasp of the meaning of a proposition in order to be experienced, and grasping a proposition requires language. This rules out pre-linguistic infants and non-human animals from experiencing emotions, a consequence which is wrong.

Such accounts also give rather strange stories for emotional feelings. For Marks, fear is having a relevant belief and a strong desire. But what is a strong desire and how do we measure it? Usual ways in a which a desire is strong and which we use to measure its strength are in how strong the compulsion is for us to act on it, and in how difficult it is to think of other things. But it then seems that someone can have an emotion without a strong desire. For example, I can be afraid of the height of the diving board at the public swimming pool. My fear manifests as an obvious and difficult-to-ignore nervous energy but, if I desire to avoid the diving board, it is not a strong desire. This is seen in the way that I easily ignore the desire as I nervously and fearfully climb the ladder to the take the plunge; my desire does not dominate my thoughts. Perhaps my fearful desire is strong but outweighed by other desires, such as the desire to do something I have never done before or my desire to win a medal. But, again, on what basis do we say it is a strong desire? It neither dominates my thoughts nor noticeably compels me to act, affecting my subsequent action choices. It is not identifiable in a way in which a strong desire is normally identifiable. If my desire is merely a weak one, then there is nothing special to the belief-desire set which distinguishes the same beliefs and desires but without the emotion.

Finally, we might question whether emotions are the kind of thing which can be broken down into beliefs and desires in the first place. A range of philosophers including Peter Goldie (2000), Sabine Döring (2003), Bennett Helm (2001) and Andrea Scarantino (2010) have all argued that emotions are neither belief-like nor desire-like because they do not have the correct directions of fit. Beliefs are taken to have mind-to-world direction of fit because we revise our beliefs in light of the way things are in the world, whereas desires are taken to have a world-to-mind direction of fit because we strive to change things to match what we desire. If being a conjunction entails that we can separate out the two conjuncts, as Scarantino (2010, 361-62) argues, then if an emotion is a conjunct of belief and desire, we should be able to separate out the belief and the desire of an emotion. However, we cannot because of the very nature of the intentional content of an emotion. For example, while fear is a response to *something dangerous that obtains*, and in that way is belief-like, it is also desire-like because we are motivated to change things because there is *something dangerous that obtains*. As Scarantino puts it:

> . . . being motivated to get away from a snake while fearing it is inseparable from representing it as dangerous and not freely employable in practical inferences, as the desire to get away from a snake would be. (Scarantino 2010, 361)

I have only given a very brief overview of belief-desire models and the kind of problems they face because such models are widely agreed to be unsuccessful. They both over-intellectualise emotions and are not able to give a correct characterisation of the kind of state that emotions are. We should therefore turn to more credible contenders for cognitive accounts of emotions.

4.2.2 JUDGEMENT ACCOUNTS

Another kind of cognitive approach to emotions, one that is aimed at getting around the over-intellectualisation challenge, is that emotions are or involve judgements of a particular kind.

Judgements are usually taken to involve taking a proposition to be true. One objection to very basic forms of judgement accounts is that such accounts are no better than the belief-desire model. This is because making a judgement is forming a propositional attitude and so, like belief-desire models, this rules out animals and infants from experiencing emotions. Another objection is that it seems that we can experience an emotion but also judge that things are not as the emotion represents them. If so, then such accounts make implausible claims about the irrationality of emotional experiences (Deigh 2010, 26-27). These are both compelling objections, and so I jump directly to judgement accounts which use a wider notion of judgement in order to avoid such problems. I shall examine two versions, those of Martha Nussbaum (2001) and Robert Solomon (1976, 2004). I first introduce each of the versions, showing how they attempt to get around the problems facing the belief model, before raising two problems that affect any version of a judgement account.

Nussbaum (2001) presents a cognitive theory and even denies the importance of bodily feelings for emotions. For her, an emotion is an assent to an appearance of how our goals are faring and how we are flourishing. This assent to an appearance is a judgement and only requires receiving and processing information together with some basic positional and causal thinking. It does not require propositional thinking. In other words, the kind of judgement involved in emotions is not one which necessarily takes a proposition as its object. Because Nussbaum has this minimal cognitive requirement and does not require judging that a proposition is true, she believes that her theory can accommodate infant and animal emotions. And

because the kind of judgement she requires is not assenting to a proposition, she believes that we can also hold beliefs that are in contradiction to what our emotions are about. She dismisses the importance of bodily feelings by arguing that the phenomenology of an emotion just is the phenomenology of fully assenting to an appearance.

Solomon (1976, 2004) argues that emotions are complexes of judgements of a very particular kind. They are evaluative judgements which are subjective engagements with the world. They need not be the singular judgements we use to define them and they need not be deliberative or reflective, articulate or fully conscious. This keeps the cognitive involvement to a minimum: emotions are responses which require only recognition and we do not need to be able to articulate what we recognise (Solomon 2004, 79).

In his later work, Solomon incorporates the distinctive phenomenology of emotion into his account by arguing that it is the bodily feeling. He accepts that we must incorporate into an account of emotions the workings of the autonomic nervous system, as well as the phenomenological manifestations of bodily preparation such as assuming a response posture. In this way, his account is importantly different to Nussbaum's. While he accepts these physical aspects to an emotion, he maintains that 'these are not just sensations or perceptions of goings-on in the body' (Solomon 2004, 86)—contra the kind of feeling theories I have already discussed. Rather, '[b]oth arousal and action readiness should be subsumed under the more general phenomenological rubric of *getting engaged with the world*' (86). In other words, the bodily feelings are feelings of the body being engaged with the world, which is what an emotional judgement is for Solomon. So, the feeling of an emotion is a judgement of the body (87).

By fleshing out the relevant judgement as a kind which need not be deliberative, articulate, fully conscious, and which requires only recognition, Solomon provides a way around the objection of over-intellectualisation facing the belief-desire theorist. The emotional subject need not form explicit beliefs about the way things are; rather, she simply recognises that things are that way and the emotion is the recognition. This recognition can be non-propositional and the kind of judgement involved with emotion is 'engagement with the world', allowing for a kind of bodily judgement and thus incorporating feelings directly.

In these ways, both Solomon and Nussbaum propose accounts that aim to get around the propositional problem facing the belief theorist, and they both give some minimalist story for emotional feeling. Despite this, their accounts are not satisfactory. There are two main problems which together make up what has been called

the 'fear of flying' objection (de Sousa 2014): the judgement is not sufficient for an emotion and neither is the judgement necessary for an emotion.[11]

The first problem is that a judgement is not sufficient for an emotion. One can make an evaluative judgement of the appropriate kind yet not experience an emotion. For example, if you cut me off in traffic, I can judge that what you did was offensive and insulting, yet not feel any anger (Robinson 2004). Therefore, the evaluative judgement is not sufficient for an emotion.

The earlier Solomon attempts to get around this kind of problem by describing emotions as a particular kind of self-involved and intense judgement. But, as Jenefer Robinson (2004) points out, judgements are not usually thought of as lying on an intensity scale. Unless we do already accept that emotions are examples of such judgements, there is no reason to suppose that they are intense judgements at all. Early Solomon risks begging the question.

In later Solomon, we find a better answer. There, as we have seen, he argues that emotions are complexes of judgements, including bodily engagements with the world. While we might define anger as the judgement that what you did was offensive and insulting, he argues that it does not follow that this judgement is the judgement in an emotion, or that it is the only judgement involved. Indeed, the very fact that I am experiencing an emotion suggests otherwise. In particular there is a judgement of the body as well as other judgements which are engagements with the world.

In support of this multi-judgement view, Solomon argues that emotions are not simple one-valence things. That is, love is not always only good or pleasant; anger is not always only bad and unpleasant. They are made up of many different elements and it is simplistic to reduce them to a single judgement with a single evaluation corresponding to a single, simple, valence. This is quite true. The feelings of emotions are complex—just think of the enjoyment of self-righteous anger—and introducing a range of judgements is one way to deal with the complexity. However, Solomon's solution brings with it new worries which undermine the idea that emotions are judgements. I shall return to these below, once I have presented Nussbaum's solution to the problem.

Nussbuam also denies that the judgements are the same in the case of an emotion and in the case of an analogous judgement without the emotion. If an emotional judgement is an assent to an appearance and if I judge that what you did was of-

[11]This objection is the 'fear of flying' objection because of the example that Michael Stocker (1992) uses: I may be aware that flying is the best means of transport and judge that it is not dangerous and yet still feel afraid.

fensive and insulting but am not angry, I have not fully assented to the appearance as offensive and insulting. I may have endorsed the proposition that it is offensive and insulting, but this need not be the same as the emotional judgement. In other words, I am not making the same judgement and it is not a problem that I do not experience an emotion.

These kind of responses are dogmatic. Both Nussbaum and Solomon stipulate what a judgement is in order to incorporate emotions under the term, but are not convincing in showing why we should adopt their definitions if we did not already accept that emotions are an example, or involve examples, of a special kind of judgement. Scarantino (2010) calls this strategy of modifying and stretching the notion of judgement to fit emotions the 'Elastic Strategy'. As he forcefully argues, such a strategy makes the judgementalist model at best trivially true because, by definition, an emotion will be a judgement. But, at worst, the judgementalist model is not falsifiable because each counterexample is addressed by stretching the notion of judgement a little bit more.

With regards to Solomon's proposal, we see the Elastic Strategy at work in the way it remains mysterious how a bodily judgement integrates with what we usually think of as judgements to create a unified emotional experience. While emotions do not have simple valences, we still have an experience of anger, say, and not an experience of bodily engagement with the world plus my (negative) judgement that I have been offended plus my (positive) judgement that I am in the right. Having this multiplicity of judgements in what is, phenomenologically, a single experience is mysterious. We can adjust what we mean by 'judgement' until we no longer mean what we usually do when we talk about judging that thus-and-so is the case, which is what Solomon does. But we have no reason to thus widen our understanding of judgement in order to incorporate emotions unless we already believe that emotions are like judgements.

With regards to Nussbaum's proposal, we are left wondering why fully assenting to appearance is importantly different to neutrally judging that the appearance is the case, or why it has its own specific phenomenology. If fully assenting to an appearance feels a certain way, does that mean that people making a judgement with a propositional content which expresses that appearance are not fully assenting to how things are? This is despite believing that judgements of the usual kind are assenting to the way things are by taking an attitude of truth towards its content. For instance, a judge assents to the judgement that there has been a murder but does not necessarily feel angry. Does this mean that we are engaged in an elaborate error theory when we think that judgements are truth attitudes? These are just questions,

not arguments. I merely wish to draw attention to how much such judgement theories require adopting a very different notion of judgement which, if we did not already think that emotions involve judgements, there are no good reasons to accept and perhaps even good reasons for rejecting.

The second problem is that the judgement is not necessary for an emotion. In a famous debate in the 1980s with the cognitive psychologist Richard Lazarus, the non-cognitive psychologist Robert Zajonc argued that emotions do not involve cognitions like judgements. For example, in 'Feeling and Thinking: Preferences need no inferences', Zajonc (1980) argues that studies on recognition support the idea that faster affective reactions, such as liking something with which one is familiar, occur before the slower cognitive judgement, such as recognising that something is familiar.[12] He interpreted data like this as showing that the cognitive act takes too long for there to be an emotional experience.

For Zajonc, cognitions are full-blown judgements that happen after a sensory input, and it is clear that we can experience an affective state without making such a judgement in this sense. Therefore, on that interpretation, the judgement model must be rejected. However, as we have seen, neither Nussbaum nor Solomon think of judgement in this way and neither did Lazarus, who endorsed a broader definition of cognition. If we water down the cognitivist claims, then we could argue that the judgement is still necessary for there to be an *emotion* and not just an *affective reaction* (Barrett et al. 2007). If we follow this route, then Zajonc's data can be interpreted to show that we can have faster affective reactions, but that we only experience an emotion once some simple cognition gets involved. Barrett et al., for instance, propose a primitive kind of evaluation, one which is subconscious and functions independently of more complex cognitions. If we insist on calling such cognitions 'judgements', however, we return to the first problem of the Elastic Strategy.

Judgement models of emotions are not satisfactory. While they are better placed than belief-desire models because they do not require that emotions are fundamentally about propositions, they are ultimately trivial and uninformative. Nevertheless, the failure of judgement models does not show that all cognitive accounts of emotion are lost. It may still be the case that emotions involve a primitive kind of evaluation that is not a judgement. In the last section of this chapter, I shall explore how we can make sense of this idea by adopting a perceptual analogy.

[12]For an overview of the debate, see Schorr (2001). The debate between Lazarus and Zajonc ultimately arose out of different definitions of cognition (Leventhal and Scherer 1987) and was even about quite different features of emotion (Lewis 2010).

4.3 WEAK COGNITIVE ACCOUNTS

In section 4.1, I argued that pure non-cognitive accounts of emotion fail because they cannot accommodate the intentional element of emotion. Some cognitive involvement is required. In section 4.2, I argued that strong cognitive accounts of emotion will also not work, either because emotions come out as too cognitively complex or because the account becomes trivial. Nevertheless, emotions are still intentional and we need an account of emotion which is able to capture this.

In this last section, I explore a final option, that of weak cognitive accounts. I first introduce Scarantino's distinctions between different forms of cognitivism as a means to establish what the different kinds of claims a cognitivist could be making are. Doing so allows us to see that emotions do not have to be analysed in terms of cognitive states with which we are familiar, such as beliefs and judgements, in order to have the intentional nature that they have, and it opens the way for a whole new approach to emotions. I shall end the chapter by introducing one such approach, that of modelling emotion on perception.

4.3.1 DIFFERENT KINDS OF COGNITIVISM

In discussing cognitive accounts of emotion, Scarantino (2010) distinguishes between three different kinds of cognitive claims. His discussion helps to set up a framework to compare different cognitive accounts, to pinpoint where accounts like Nussbaum's and Solomon's go wrong, and to establish what a successful account needs to do. I shall therefore draw on it in order to motivate the kind of approach that should be taken in giving an account of occurrent emotions.

Scarantino divides the various cognitivist positions into three groups, each making one of three different claims. These groups and claims are:

1. Constitutive Cognitivism: emotions are essentially judgements.

2. Etiological Cognitivism: emotions are caused by an appraisal.

3. Representational Cognitivism: emotions represent their formal objects.

Nussbaum and Solomon are examples of the Constitutive Cognitivist because they claim that emotions are essentially judgements of a particular kind, and so experiencing an emotion is making a judgement. Constitutive Cognitivism fails for the reasons presented in the previous section. It is at best trivially true and, as Scarantino writes:

> Introducing a seemingly unitary construct into a mongrel of evaluative, physiological, phenomenological, expressive, behavioural and mental components leads us to lose track of the way in which components contribute towards something being an emotion *E*. (Scarantino 2010, 749)

Etiological Cognitivism is the subject of the Lazarus-Zajonc debate and Schachter and Singer would also fall under Etiological Cognitivism. They are all concerned with whether an emotion is caused by an appraisal or a cognition and do not necessarily say anything about whether that appraisal or cognition also constitutes the emotion. Whether or not an emotion is caused by an appraisal or a cognition of some sort will depend on what that appraisal or cognition involves. If it is a full-blown judgement, then Etiological Cognitivism is false. If it is some primitive evaluation of the kind that Barrett et al. (2007) endorse, then the jury is still out.

Representational Cognitivism only 'states that emotions, whatever they are and however they are caused, are *intentional states* of a particular sort' (Scarantino 2010, 756). An emotion could be intentional by being a judgement or belief, but it could also be intentional by being an evaluative feeling (Helm 2009), a feeling towards (Goldie 2000), pleasant or unpleasant feelings about some features attributed to the object of an emotion (Greenspan 1988), an evaluative construal of how things are (Roberts 1988), perhaps an evaluative attitude (Deonna and Teroni 2012), or even a perception of an evaluative property—the last of which I shall use as an illustration of a weaker cognitive account in section 4.3.2. Despite the problems with the cognitive accounts I have discussed, by distinguishing between Representational and Constitutive Cognitivism, we are not forced into abandoning cognitivism *tout court*.

The insightful idea behind both Solomon's and Nussbaum's accounts is that we are recognising a situation as being a certain way, but we now see that the kind of recognition need not be judging that it is that way and we can conclude that, just in virtue of having an intentional nature, emotions do not have to be analysed in terms of cognitive states with which we are familiar, such as beliefs and judgements. In fact, we cannot do so.

4.3.2 Perceptual models

In this last section, I introduce one example of a way in which an emotion could be intentional without involving a belief or a judgement, and that is by being like perception. I focus on perceptual accounts to illustrate a weaker cognitive account rather than one of the others I named above for three main reasons.

The first reason is that I want to illustrate how an intentional state can be minimally cognitive but have a rich intentional nature. The perceptual model illustrates this well. The second reason is that many of the weaker accounts can themselves either be interpreted in ways that are compatible with a perceptual reading—for example, the emotional 'modality', as such, could be an evaluative feeling or a feeling towards—or are not substantial improvements on the judgement accounts.[13] Further, as John Deigh (2010) and Michael Brady (2013) both draw attention to, the shift within weaker cognitive accounts is towards accounts that are perceptual in nature. The third reason is that I wish to stay as neutral as possible with regards to which account we should adopt, and my discussion of the perceptual model will be general in a way that discussions of an individual proposal cannot. Ultimately, my aim in focusing on the perceptual model is to provide an illustrative account of emotions which I can use to structure my subsequent discussions, one which is at least, if not more, plausible than its contenders. The most important commitment I have is to the way in which, on the perceptual model, there is a robust sense in which emotions inform us of how things are.

With this in mind, I will begin by laying out what the perceptual model generally endorses, showing how, as an analogy, it illustrates a way in which emotions can be intentional states but weakly cognitive. I will emphasise how emotions can inform us of how things are, something which the perceptual model highlights nicely. I will end the section by flagging a few worries but concluding that, nonetheless, the perceptual model is a good working model to have to hand.

While the details of the different perceptual accounts vary,[14] the core idea is that emotions are (or are like) perceptual experiences.[15] As we saw with Prinz (2004), Prinz thinks that emotions are perceptual experiences of both bodily changes and of the formal object of an emotion. The dominant view amongst perceptual theorists, however, is that emotions are perceptual experiences of values or evaluative prop-

[13]Deonna and Teroni (2012, 55) argue that accounts like Greenspan's and Roberts' are still too cognitively complex and Brady (2013, 28) argues that such accounts make emotions active in a way that they are not.

[14]Versions of the perceptual model or models which have similar core claims can be found in de Sousa (1987), Jones (2003), Deonna (2006), Döring (2007, 2008), Tappolet (2005, 2012). Critics of different aspects of the perceptual model include Deigh (2010), Salmela (2011), Whiting (2012), Brady (2013).

[15]We need to be aware of the difference between 'perception' and 'perceptual experience'. As Brady (2013, 46-47) notes, perceptions are factive but emotions are not. Perceptual experiences, on the other hand, are not factive. However, emotion theorists do not seem to use 'perception' in the same factive way as perceptual theorists, and we should not interpret the emotion theorist as necessarily endorsing a view of emotions as factive. We should therefore rather compare emotions with perceptual experiences.

erties; that is, they are perceptual experiences of the formal object of an emotion. Christine Tappolet writes:

> According to the perceptual account, emotions are perceptions of values. Thus, fear would consist in perceiving something as fearsome, disgust as perceiving something as disgusting, shame in perceiving something as shameful, and so forth for every distinct kind of emotion. (Tappolet 2012, 206-207)

We could interpret the perceptual claim in at least three ways, two literal and one as an analogy. On the first literal interpretation, emotions are their own form of sensory modality. One could say that the modalities include seeing, hearing, tasting, touching, proprioception—and emoting. Prinz (2004), for example, argues that emotion is 'perceiving one's place in the world' via perceiving bodily states. Through a kind of bodily feeling, emotions are a special kind of perception. He does not think that the dependence that emotion has on other states undermines the thought that it is a distinct modality; he in fact argues that other modalities can also be dependent on other modalities, as happens with synaesthesia.

This interpretation is problematic. Demian Whiting (2012) argues that the feeling of an emotion, such as the edginess of fear, does not alone represent the evaluative property, such as the dangerous or the frightening. He writes, 'the representation of *there being a dangerous (or fearsome) object in front of me* is no part of the experience of the unpleasant edgy sensation that pervades my guts and limbs when I am frightened' (Whiting 2012, 97).[16] This I agree with, as we saw in section 4.1. If the emotional modality does consist of feelings, then in order also to be emotional feelings they have to incorporate inputs from sources of cognition. In general, emotions depend on other modalities as well as other brain states and processes, what we might call their 'cognitive bases' (Deonna and Teroni 2012, 5). Or, as Fabrice Teroni puts it, 'emotions are not independent modes of access to particular objects' (Teroni 2007, 404). I am therefore not inclined to adopt this interpretation of the perceptual model.

On the second literal interpretation, emotions piggy-back on our established senses, so to speak. Experiencing fear would be the same as seeing or hearing something as dangerous. However, emotion is not linked to any particular perceptual modality and, while seeing or hearing might be the source or the cognitive base of an emotion, they are not the same as the emotion (Whiting 2012, 95-96).

[16] Salmela (2011) presses a similar objection to a literal interpretation.

We are left with a non-literal interpretation. On such an interpretation, we can understand emotional experience as being analogous to or generally like perceptual experience in that we experience, emotionally, some evaluative property just like we experience, visually, some visual property. For example, a visual experience of redness is an experience of the property of red. An experience of fear, then, is an experience of danger. This non-literal and very loose interpretation is enough to have some model of how an intentional state which is not a belief or judgement can still be intentional.

There are three reasons why the perceptual model, even on this non-literal and minimal interpretation, is potentially fruitful in light of my argument to this point. The first is that perception provides a model of an intentional state which does not have propositions as its object like a belief or judgement does. Nevertheless, its content is expressible propositionally and we can have cognitively rich perceptual experiences. It therefore avoids the problem facing belief accounts and strong judgement accounts.

It is beyond my current scope to engage in debates in the philosophy of perception about what perception and the content of perceptual experiences are. Advocates of the perceptual model of emotion tend to work within the representationalist approach to perception according to which perceptual experiences have representational content. Many representationalists with regards to perception, however, conclude that the content of perception is propositional (Tye 1995; Byrne 2001). But this is by no mains a required tenet. For example, Tim Crane (2009) rejects that the content of perception must be propositional even if it is representational. We may think that the content of perception is propositional because it has accuracy conditions and can be true or false, and truth and falsity are properties of propositions. However, as Crane argues, a picture is a representation of a state of affairs which is also assessable for accuracy, but a picture is not a proposition that can be true or false even if it can be described in propositional terms. Therefore, something can be a representation that is assessable for accuracy without being a proposition. For the purposes of the perceptual model, it is important to understand perceptual content as non-propositional, even if the content is the kind of thing that can be expressed propositionally and assessed for accuracy.[17]

The second reason why the perceptual model is fruitful is that it provides a model of an intentional state which is not a truth attitude like belief or judgement. It

[17]A number of emotion theorists defer to Crane, both for his arguments about the non-conceptual content of experience as well as for his understanding of intentionality. These include Goldie (2000), Döring (2007) and Salmela (2011).

therefore creates room for the possibility of holding beliefs and making judgements which conflict with our emotions without being found undesirably irrational all too often.

The third reason is that the perceptual model draws attention to the way in which emotions can inform us about the world. As Brady writes, it is 'a platitude that emotions constitute reactions to objects, events, and states of affairs that are potentially significant or important to us' (Brady 2013, 10). From our own experiences of emotions, we see the truth of such a platitude. For instance, when in a dangerous situation you react in fear; or when around your children you react with loving feelings. Through experiences like these, we learn that emotions are reactions to things that are of significance to us. But many research programmes in psychology are built on and provide evidence of the way in which we react emotionally to things of potential significance. For example, emotions are understood as being 'affect programmes', which are short-term reactions to a certain class of inputs (see, for example, Ekman 1977; Griffiths 1997), or as involving action tendencies which prime us for suitable responses within the environment (see, for example, Frijda 1986), or as involving appraisals (see, for example, Schorr 2001; Moors 2014).

If emotions are reactions like this, then another platitude becomes apparent, which is that, as Karen Jones writes, '[e]mpirical work on and common observation of the emotions tells us that our emotions sometimes key us to the presence of real and important reason-giving considerations' (Jones 2003, 181). We can and do learn about the world through our emotions. For example, if you feel intense anger at the acquittal of some apparent felon, you can gain self-knowledge by learning that you cared about the case a lot more than you thought you did, but you also learn something about the world, namely that something you care about has been infringed on, or even that there has been an offence. Or, walking down a dark street, you might feel a twinge of fear which makes you reflect on where you are and realise that you are actually in danger. Through the emotional experience, you come to learn something about the situation that is of potential significance.[18]

In Chapter 3, I proposed that, if emotions are passive like they have been historically treated, then they could give us information about the things which 'act' on us, much like perception can. And much like through perception we can come to know about things and whether they are present, through experiencing an emotion we can come to know about how things are and their significance. The perceptual model provides a way in which to understand how emotions are reactions to and

[18]See Brady (2013, 10-25) for a more detailed overview of the epistemic value emotions have, with a particular focus on how they direct our attention to salient features of the world.

inform us about the world.

The perceptual model therefore provides a neat illustration of how emotions can be intentional and how they can give us information of how things are. There are, of course, problems facing the perceptual model. One of these, argued by Deigh and pertinent to my current discussion, is that perceiving an evaluative property, namely the formal object, still requires application of the concept of the formal object. For instance, perceiving something as dangerous requires the application of the concept of danger. If so, then emotions potentially remain too cognitive for pre-linguistic infants and non-human animals (Deigh 2010, 29-32).

I do not think that this poses a problem for the perceptual theorist. Consider the example that Deigh raises about a dog fearing a man with a stick. Deigh proposes that the dog does not see the man as *dangerous*, but rather sees the man as an imminent source of pain. This is owing to previous experiences of the man's threatening behaviour leading to his (the dog's) being struck, combined with the perception of the man in a similar state. We should, however, keep distinct our *concept* of the formal object from the actual property, and doing so avoids Deigh's objection. We *describe* the formal object of fear as the dangerous or the frightening. Being dangerous, in contrast, requires that there is a threatening relation that holds between the subject and some object in her environment and we can experience that relation without applying the description. The dog thus does perceive something dangerous: it perceives the man as an imminent source of pain, where being a source of pain is standing in a harmful relation with himself. As creatures get more cognitively complex, they may start to recognise the formal object in more complex scenarios and they may start to recognise danger *as danger*, but it is possible to see some evaluative relation as holding without applying a concept. I shall say more about the concept of the formal object in Chapter 6.

Other problems with the perceptual model arise specific to not adopting a literal interpretation. In particular, there may be important limits to how far the model can go or important ways in which emotion and perception are disanalogous. For example, emotions appear to have different epistemic roles to perception (Wild 2008; Salmela 2011; Brady 2013), and we might worry about the kind of evaluative information which emotions are supposed to be perceptions of as being disanalogous to the kind of information that normal perception gives us (Salmela 2011). I shall address these kinds of problems in Chapter 7.

The way in which I develop the concept of the formal object in Chapter 6 and the way in which I respond to these objections in Chapter 7 may take us further away from a perceptual analogy. My current purpose, however, has been not to

argue for the perceptual model rather than to give an illustration of how we could understand emotion as an intentional state that is minimally cognitive but which still gives us information about how things are, thereby avoiding the problems facing both non-cognitive and cognitive accounts. My core motivations for introducing the perceptual model remain: it illustrates an understanding of emotions as intentional and it draws attention to the informative role emotions can have. All accounts face problems so one has to choose what to use as a working model, and the perceptual model seems at least as good, but perhaps even one of the best, models to use.

4.4 Conclusion

In this chapter, I have had fairly limited aims. I have not set out to provide a comprehensive analysis of what emotions are, nor even to whole-heartedly endorse any one account. Rather, I have argued two things. First, I have argued that emotions are about something other than our bodies. Non-cognitive accounts of emotions are unable to accommodate this insight satisfactorily without bringing in some kind of cognitive involvement. Second, I have argued that the cognitive involvement is not that of a belief or a judgement. Rather, emotions are another kind of intentional state which gives us information about how things are, beyond our bodies.

I ended by proposing the perceptual model as an account of emotions which illustrates how emotions can be intentional but minimally cognitive. My arguments about whether emotions provide access to reasons for acting will not turn on whether or not the perceptual model is in fact correct. What they will turn on is whether emotions give us information which could be the agent's reason for acting, either through being like perception or involving some other kind of cognition. And since the perceptual model can accommodate this fact and is at least as plausible as the other contenders, I shall adopt it as a working framework.

As I noted right at the outset, there are two corollaries of the discussion in this chapter which are important for the rest of my thesis. First, if emotions are about things other than our bodies, then they give us information of something other than our bodies. This information could be a fact in light of which we act. I shall explore this in Chapter 6. Second, if some kind of weak cognitive involvement is necessary for emotions and that need not be a belief or a judgement, then emotions could play a role in action which is different to that of a belief or judgement. I shall expand on how this is possible in Chapter 7.

Before getting there, in Chapter 5, I address some more important background and lay out what acting in light of a reason and rationalising explanations are.

CHAPTER 5

RATIONALISING EXPLANATIONS

In chapters 2 and 3, I introduced examples of explanations of action in terms of an emotion and tentatively proposed that some of those explanations could involve an emotion in virtue of its intentional nature, where acting in light of an emotion could be acting for a reason in some way. The emotions in Category D, for instance, are the agent's reasons because they are states which the agent acts to address. But perhaps the emotions in Category E, where the agent acts because of the emotion but not in order to address it, also provide the agent's reasons through being about or elicited by those reasons in the right kind of way. In Chapter 4, I argued that emotions are best understood as intentional states that are weakly cognitive but which provide us with information about how things are. The information they provide could be those reasons in light of which we act.

In order to establish whether emotions can play this kind of reason-providing role in action and action explanations, it is necessary to look closer at what the information is and at whether emotions provide these reasons in the right kind of way. Before being able to do so, however, it is necessary to clarify what reasons are, what rational action is, and what a rationalising explanation would involve. This is my task for the current chapter. Ultimately, I shall argue that rational action is action done in light of a reason and side with those who argue that a reason is a fact, that a rationalising explanation gives the agent's reason for acting, and that even if a rationalising explanation is not available, we can still give an explanation showing how an action was reasonable. The topics of rationalising explanations and rational action are large, controversial and unwieldy. The positions I adopt form a plausible account albeit not free from their own difficulties—but none is—and, while I cannot here provide a comprehensive defence of them, I shall motivate why we should adopt them.

In section 5.1, I begin by identifying three necessary conditions for rational

action, including the condition that the agent acts in light of a reason in order to achieve something. This will set me up to turn to explanations of rational action in the rest of the chapter.

In section 5.2, I introduce Donald Davidson's influential thesis that a rationalising explanation must be in terms of a belief-desire combination. Even if this is the case, something I will dispute, I argue that it does not follow that the agent's motivating reason—the reason for which she acts—is also a psychological state like a belief or desire. Instead, I side with those who have argued that a motivating reason is a fact that counts in favour of the action and which can be stated by a true proposition. I reject the view that a rationalising explanation must be in terms of a state of belief, a state of desire or a combination, or that such an explanation is the central kind of rationalising explanation. This is because what makes the action rational is that the agent has a motivating reason. An explanation of rational action *qua* rational which gives the agent's motivating reason will be the central kind of rationalising explanation and such an explanation need not make reference to psychological states like beliefs and desires. I shall end the section with a problem for the proposal that the central kind of rationalising explanations are in terms of motivating reasons. This is the problem that not all action can be explained in terms of facts. For example, an agent who is wrong about the facts, on the view I accept, does not have a motivating reason; nevertheless, it seems that we can still explain her action as rational by identifying her states of belief and desire. Because it seems that we can explain all action in terms of states of belief and desire, in contrast, it seems that the central kind of rationalising explanation is not one in terms of motivating reasons as facts or, the even stronger conclusion, that motivating reasons are not what rationalise the action. I consider one kind of response to the problem, which is to claim that we can explain all action in terms of motivating reasons, but conclude that such a response does not work.

In sections 5.3 and 5.4, I take a different tactic to argue that this problem does not show that motivating reasons do not have a central place in rationalising action. In section 5.3, I argue that not all action can be explained in terms of states of belief and desire. I then propose an alternative model of explanation. On this model, we can give a rationalising explanation in terms of the agent's motivating reason, but we can also give an explanation in terms of the agent's states of belief and desire or another psychological state. The former explains by giving the motivating reason and provides a rationalising explanation; the latter explains by giving an explanatory reason. Sometimes motivating and explanatory reasons will overlap, but they need not. I shall end by discussing how the latter kind of explanation may show that the

action was reasonable, but it does not show that it was rational or done for a reason.

By the end of the chapter, I will have introduced a framework of explanation which helps to formally describe the difference between the different categories of emotion explanation I identified in Chapter 2, as well as a framework which makes genuine the questions of whether we can act in light of a reason when acting on the basis of an emotional experience and whether emotions can provide a kind of rationalising explanation. I will conclude the chapter by describing the categories of explanation in light of the framework and highlighting how the framework opens up the rationalising potential of emotions. In particular, if rationalising explanations must be articulated in terms of states of belief and desire alone and if emotions are not reducible to beliefs and desires, as I argued in Chapter 4, then an explanation citing an emotion would not be a form of rationalising explanation. However, if explanations can be in terms of motivating reasons and if emotions could provide those motivating reasons, then giving an emotion in an explanation could be a form of showing that the agent acted in light of a reason and a form of explaining the action as reasonable.

5.1 Rational action

In this section, I identify three necessary conditions for rational action. These are that the agent controls the action in some way, that she is aware of the action under a certain description, and that she acts in light of a reason in order to achieve something.[1]

What would make an action a rational one?[2] We could begin by saying that rational action is something a rational agent does.[3] This, however, is too inclusive: it includes things that a rational agent does but which we do not think of as actions, let alone rational. For example, I am a rational agent. Two things I do are breathe and digest my food; if I did not do either, I would no longer be a rational agent as I would not be living. If I am pushed and stumble forward, stumbling forward is something that I, a rational agent, do. But none of these is rational let alone an action. What is the difference between things like breathing, digesting and stumbling on the one hand, and what we would want to call rational action on the other?

[1] I draw largely on Joseph Raz (1999) and Maria Alvarez (2009) for these conditions.

[2] It may be the case that all action is rational action but I shall not assume this.

[3] Defining what is a rational agent is itself a difficult task. An agent, quite generally, is something that effects change. A rational agent, at the very least, effects change through a rational capacity, such as by being capable of deliberating and weighing up choices and acting on those choices. Humans are the paradigmatic rational agent.

One thing that is lacking in the cases of breathing, digesting and stumbling is control. As Joseph Raz (1999, 39-40) discusses, the kind of control that is necessary for rational action comes in different degrees and can be control over initiation of the action, including being able to stop or avoid an action, or control over the execution of an action. Digesting my food is something over which I have no control; therefore, it does not meet the control requirement for rational action. If I am pushed and stumble forward, I do not have control over the initiation of what I do. If I were to continue forward and playfully run into my friend, then I do have control over the execution because I would have been able to stop earlier than I did. While my initial stumbling was not an action, my playful continuation is. While I do not have full control over whether I breathe or not, I can sometimes control the execution of my breathing—such as holding my breath on entering a stinky tube carriage, taking a large breath in order to play saxophone, or controlling my breathing pattern while meditating. When I control my breathing in these ways, my breathing plausibly becomes an action, distinct from my automatic breathing. Having control over what is being done is thus the first condition for rational action.[4]

Control, however, is not enough. Suppose that, when inhaling deeply in preparation of playing a long passage on the saxophone, I also unknowingly swallow a fly. I control my inhalation and my action is of inhaling; it is not of swallowing a fly even though this is something I also do. What is the difference between these two descriptions of what I am doing? The difference is that, in the former, I am aware of what I am doing as inhaling but, in the latter, I am not aware of what I am doing as swallowing a fly. We therefore have a second condition for rational action: being aware of what one is doing.

Awareness and control, alone, are still not enough for rational action. Consider another case of non-rational action, this time of an animal. An otter controls its breath as it dives under water and it is surely aware that it is doing so. The action is not a rational action, however. In order to be a rational action, the agent needs to be acting purposively and for a reason.

The otter, in one sense, does act for a reason. It holds its breath in order to dive underwater without its lungs filling with water, and in this sense does act purposively and for a reason. It does not, however, 'act for a reason' in another sense of the phrase.

[4]Exactly what kind of control is required, as Alvarez (2009, 294) draws attention to, is controversial. The relevant kind of control might include that one would have been able to refrain from doing what one did, but this is a hotly debated topic. I do not need to say anything about these debates; the point is simply that some kind of control is important.

Acting for a reason, in this sense, is acting in light of some consideration and in pursuit of a goal, where the consideration counts in favour of acting and the agent is aware of the relationship between the consideration, the action, and the goal. The considerations which count in favour of acting are *the agent's reason* and not just the reason why the agent acts. It is this sense of acting for a reason which is crucial for rational action and I shall refer to it as 'acting *in light of* a reason'. As we shall see below, some philosophers have argued that the considerations in light of which we act are our states of believing and desiring; others have argued that the considerations in light of which we act are facts or propositions. Either way, what sets the otter's breath control apart from my breath control is that I am acting in light of some considerations in order to achieve something and am aware of the relationship. Thus, in addition to having control and being aware of what one is doing, rational action requires a third condition: the agent is acting in light of a reason in order to achieve something.

Rational action, on this understanding, is acting in light of a reason. As we shall see, we must distinguish between rational action and intentional action, where the latter is action under our control and with awareness, and is not done accidentally but not necessarily done in light of a reason.[5] Whether intentional action is always done for a reason is a debated topic and I shall return to it in section 5.3. It is enough for my current discussion that, as far as terminology goes, 'rational' action is accepted to be action in light of a reason, regardless of whether it is the same as intentional action. Action not done in light of a reason need not be irrational, just arational.

We now have three conditions for rational action: the agent is in control of the action in some way and to some degree, she is aware of what she is doing, and she acts in light of a reason in order to achieve something. As rational agents, we act with purpose and in order to achieve our aims, and we act for reasons which will help us to do that. Acting in light of a reason is what sets rational action apart from other action.

If these are three conditions for rational action and if acting in light of a reason is the characteristic condition, then an explanation of rational action *qua* rational action will show how an action is something the agent did in light of a reason.

[5] I draw on Alvarez (2009) for these requirements for intentional action.

5.2 BELIEFS, DESIRES AND MOTIVATING REASONS

The aim of explaining an action is to come to know why the agent did what she did, and the aim of explaining rational action *qua* rational is to show that the agent acted in light of a reason and come to know the reason in light of which the agent acted. This kind of explanation is a rationalising explanation. The agent's reasons for doing what she did are often called the agent's motivating reasons (see Dancy 2000; Audi 2002; Alvarez 2010).[6] The questions to be investigated, then, are: (i) what are the agent's motivating reasons and (ii) how do we rationalise an action by reference to them?

Following Donald Davidson's seminal paper 'Actions, Reasons, and Causes' (1963), the dominant account of rationalising explanation of action, often called the 'standard account' (Everson 2010), holds that rational action is explained by the agent's states of belief and desire that caused the action. The standard account has three further commitments. First, all intentional action will be explicable in terms of states of belief and desire. Second, these beliefs and desires, it is thought, are the agent's reasons for the action; they are the agent's motivating reasons. Third, all intentional action is rational action, or action done for a reason.

I am not concerned with whether the agent's reasons will be the same as the cause of the action or whether the only rationalising explanation is a kind of causal explanation, two important claims to Davidson's own philosophy. We can explore the nature of rationalising explanations without taking a stand. This is because we can give explanations which increase understanding but which do not show a causal link, and so there is nothing necessary to an explanation that it gives a cause. For example, Napoleon walked under the archway without bending because he was a short man. Napoleon's height did not cause him to walk under the archway without bending, but it is still why he did it. Perhaps rationalising explanations show why an agent acted without necessarily showing that the reasons why the agent acted caused her to act. I shall thus rather put pressure on the standard account by arguing, first, that states of belief and desire are not the agent's motivating reasons and, second, that not all intentional action will be explicable in terms of states of belief and desire.

[6] I am not concerned here with normative reasons, reasons which independently count in favour of or justify an action, reasons for which an agent ought to act. My focus is on explanations of action that an agent has performed, seemingly for a reason, and so I set aside controversial debates about whether normative reasons are internal or external, and whether one can have a normative reason and not be motivated by it. I do, however, accept that normative reasons must be the kind of thing for which one could act, that is, that normative reasons can be motivating reasons and vice versa.

In the rest of this section I address the first point. In section 5.2.1, focusing on Davidson's proposal, I argue that explaining an action in terms of the agent's states of belief and desire does not entail that the agent's motivating reasons for the action are states of belief and desire. Instead, we should accept that the agent's motivating reasons are facts in light of which she sees the action as good or desirable. This implies that an agent's motivating reasons are not psychological states. Even if motivating reasons are not psychological states, it may still be the case that we can rationalise an action by giving the agent's states of belief and desire. Indeed, it might be argued that we can rationalise all action in terms of states of belief and desire, suggesting that motivating reasons do not have a central role in rationalising action. This is a problem I raise in section 5.2.2.

5.2.1 Explaining an action in terms of beliefs and desires

Davidson calls explanations which explain rational action 'rationalisations'. These are explanations where '*the reason explains* the action by giving *the agent's reason* for doing what he did' (Davidson 1963, 685, my emphasis). I shall suggest that the first reason mentioned in the quotation, the reason that explains, need not be the second reason, the agent's reason. To do so, I present Davidson's account of explanations of action in terms of the agent's beliefs and desires, then argue that it does not follow that the agent's motivating reasons are the same as the reasons which explain the action. Indeed, I argue, the agent's motivating reasons are facts in light of which the agent sees the action as a good thing to do. They are facts which count in favour of the action and which can be stated by a true proposition.

According to Davidson, not just any reason will rationalise an action:

> A reason rationalizes an action only if it leads us to see something the agent saw, or thought he saw, in his action—some feature, consequence, or aspect of the action the agent wanted, desired, prized, held dear, thought dutiful, beneficial, obligatory, or agreeable. We cannot explain why someone did what he did simply by saying the particular action appealed to him; we must indicate what it was about the action that appealed. (Davidson 1963, 685)

We see in this quotation that Davidson identifies two necessary features of a reason if it is to rationalise an action. The first is that it captures something the agent saw or thought he saw in the action. In other words, the agent believes (or knows, perceives, notices, remembers) something about some feature of the action that makes it appeal to him. The second feature is that the agent must desire, want or

have some other pro-attitude to that appealing feature which he sees in the action. A pro-attitude is just what the name suggests: an attitude of the agent in favour of action of the relevant kind. Davidson talks about 'pro-attitude' rather than just desire or want because he wants to include cases where we may not desire to do something, but still have a pro-attitude towards doing it, like having a sense of duty or an obligation.

A reason which has these two features is what Davidson calls a 'primary reason'. Specifically, the primary reason is made up of a pro-attitude towards a class of action, along with the belief that the action *under a certain description* falls into that class. By requiring a certain description, Davidson aims to rule out descriptions of the action for which the agent could not have acted. For example, I open my mouth to take a deep breath in order to blow on the saxophone and inadvertently swallow a fly. Because I did not know that there was a fly there, we cannot say that I opened my mouth in order to swallow a fly, even though I did intentionally open my mouth and inhale deeply. With the description condition, we can rule out the description of my action as 'swallowing the fly' as one done for a reason because I do not have the accompanying belief that this action is an action under the description of 'swallowing a fly', even if I were to have a pro-attitude towards such action.

We now have the outline of what a primary reason involves and can focus on some details. In particular, if there is a primary reason when the agent has a belief and a pro-attitude of the right description, then is the agent's motivating reason also made up of that belief and pro-attitude? Many philosophers have interpreted Davidson as saying that the agent's reasons for acting, her motivating reasons, are her psychological states. This, however, does not follow smoothly from Davidson's own discussion.

Going back to the indented quotation above, we see that a reason 'rationalizes an action only if it leads us to see something the agent saw, or thought he saw, in his action' and what the agent saw is 'some feature, consequence, or aspect of the action the agent wanted'. The reason rationalises by leading us to see *what* the agent saw in his action. In other words, the reason rationalises by picking out something external to the agent, albeit of which the agent has beliefs and desires. Davidson carries on to say that, 'we cannot explain why someone did what he did simply by saying the particular action appealed to him; we must indicate what it was about the action that appealed'. What does the work in showing that an action is one of an agent is not so much that the agent has states of belief and a pro-attitude, but that there is something which he believes and something which he has a pro-attitude towards. That is, what does the rationalising work is what makes the action

something to do in the first place.

Despite this, one school of thought, what can be described as the 'Humean' view,[7] proposes that the motivating reason is constituted by our state of desire. On such a view, if we have a motivating reason to do something, it is because we desire doing the action or desire its outcomes.

However, it is arguable that the reason why we desire certain things is that we see something desirable in them. To see this, consider Elizabeth Anscombe's example of a man wanting a saucer of mud (Anscombe 1957, 30). In such a case, we would try to find out *why* the man wants the saucer of mud, *why* he finds it desirable. If there is no answer other than that a saucer of mud is the object of his desire, we are dealing in nonsense. In Anscombe's evocative words, we will dismiss him as 'a dull babbling loon'. In order for a desire to make sense, it needs to have a 'desirability characterisation' which shows why it is desirable. What characterises the object of desire as desirable are the reasons for desiring it.

But if we have reasons for desiring what we desire, then those reasons are also why we act. We might say that the reasons transmit to the action (Dancy 2000, 35-38). What this means is that we act because the action is good in light of something desirable, not because we desire to act or desire the outcomes—even though we may also have these desires. For example, if I desire to play the saxophone, the reason why I desire to play is because it is enjoyable and fun. When I play the saxophone, my reason is that it is enjoyable and fun, not that I *desire* to play the saxophone or *desire* something that is enjoyable and fun.

According to another school of thought, what can be called 'psychologism' (Dancy 2000, 14), the motivating reason is constituted by some state of belief of the agent, perhaps together with a desire. Michael Smith (1994), for example, accepts that normative reasons are facts or truths which justify or speak in favour of an action—he therefore allows that what makes an action desirable is a fact—but nevertheless argues that the motivating reason is a belief-desire compound. This is because we form a belief about the normative reasons and it is this belief which causes us to desire to act. In this way, the reason in light of which the agent acts, the motivating reason, is a belief-desire compound.

We need to be wary of conflating the agent's reason with background or enabling conditions for the agent's having a reason, however. An enabling condition is a 'consideration that is required for the explanation to go through, but which is not itself a part of that explanation' (Dancy 2000, 127). Believing may be necessary for

[7]Jonathan Dancy (2000, 10), calls this school of thought the 'Humean' view, as does R. Jay Wallace (2014). I also use Dancy's name for the following view, 'psychologism'.

an agent to become aware that there is a reason to act, and it may be a necessary enabling condition for an explanation to go through, but it does not follow that it is itself her reason to act. To see this, we can actually go back to some of Davidson's later writings.

In 'Two Paradoxes of Irrationality' and 'Problems in the Explanation of Action',[8] Davidson draws attention to the fact that what is important for reasons is the kind of logical relations which stands between the *contents* of the mental states. It is not some relation that holds between different psychological states. This is because reasons play a role in practical and theoretical reasoning; and, in order to do this, they need to be the kinds of things which can stand in logical relations to each other: we make inferences on the basis of reasons, for example. So, my state of belief that the compulsory research seminar is on Wednesdays, as a state, does not stand in a logical relation to or entail other beliefs and attitudes. It does not play a role in reasoning. It is the content of my belief which stands in logical relation to other beliefs and attitudes, such as the belief that if today is Wednesday, I ought to go to the seminar. A belief-desire compound is thus the wrong kind of thing to be a reason.

A final point which brings out why the agent's motivating reason should not be identified with psychological states is that we can explain an action in terms of psychological states that are not motivating reasons. We need to be able to say which psychological states are motivating reasons and which are not, and a view identifying motivating reasons with psychological states cannot.

Consider, for example, James who does not go to the orientation week party because he is shy.[9] We could explain his action by saying that he is shy. We could also explain his action by saying that he believed that there would be many strangers there. Both explanations are in terms of his psychological states and, while the latter is an explanation in terms of what he takes to be reasons, the former is not. If the agent's reasons are psychological states, however, we cannot say which one is in terms of the agent's reasons. One might respond that the motivating reason must be a belief-desire compound; shyness is neither a belief nor a desire, therefore shyness cannot be the motivating reason. But this is jumping a step. What is it about a belief or a desire which makes it the right kind of psychological state to constitute a motivating reason? It might seem that they are the right kind of state because they are the ones in light of which we act. James sees staying away from the party as a good thing to do in light of believing that there will be many strangers there. If he

[8]Both reprinted in the collection *Problems of Rationality* (2004).

[9]This kind of example is discussed in Dancy (2004), Alvarez (2010) and Littlejohn (2012).

did not believe that there were going to be many strangers there, he would not see staying away from the party as a good thing to do (other things being equal). But this will not work: it is in light of being shy that he also sees staying away from the party a good thing to do. If he were not shy, he would not see staying away from the party as a good thing to do (other things being equal).

Perhaps his shyness is not his motivating reason because it is not his *consideration* in favour of staying away. He does not think, 'Because I am shy, I shall stay at home.' But this response does not help because the same can be said for other psychological states. James does not think, 'Because I *believe* that there will be strangers, I shall stay at home.' James' state of believing is not his consideration in favour of staying away; rather, it is the content of his belief, that there will be strangers. The psychological states, alone, are not able to distinguish between the agent's reasons and other reasons which are also psychological states.

If we accept that the content of the beliefs and desires is what makes up the agent's reason, we can make the distinction. James does not act in light of the content of shyness, whatever that would be, and so his shyness is not his reason for acting. It is the consideration about which James forms a belief which is his motivating reason, not his belief.

We therefore see that even if a primary reason is made up of how the agent sees a situation through what she believes and desires, it does not follow that the agent's own reason for acting is a state of belief and a state of desire. In fact, the agent's motivating reason will be what makes the action desirable, and a primary reason gives the agent's reason for what she did because, amongst other things, it identifies what the agent finds desirable.

The agent's reasons for acting, the motivating reasons, are best understood as 'the considerations in the light of which the agent chose to do what he did' (Dancy 2000, 6). I side with those who argue that these considerations are facts about which we can form beliefs. Motivating reasons, on this view, are facts in light of which the agent sees the action as a good thing to do. They are facts which count in favour of action and which can be stated by true propositions. This position on reasons is controversial and it is beyond the scope of my thesis to defend it more than I have done so far and will do in the rest of this chapter, but developments and defences of different versions of reasons-as-facts can be found in Raz (1986, 1999), Scanlon (1999), Alvarez (2008, 2010) and Littlejohn (2012).

5.2.2 A PROBLEM FOR MOTIVATING REASONS

It is all very well to mark a distinction between the reasons which explain and the agent's motivating reasons, but the reasons which rationalise an action may still be states of belief and desire. In this subsection, I begin by going over Smith's account that distinguishes between motivating reasons as constituted by psychological states and normative reasons as facts.[10] Because all action can be explained in terms of states of belief and desire but not in terms of facts, or so Smith claims, Smith thinks that belief-desire explanations are the central kind of rationalising explanation because they are what rationalise the action. We could reject this argument in two ways. We could argue that we can explain all action in terms of facts, in some way. Or we could deny that all action can be explained in terms of beliefs and desires. In the rest of this subsection, I show that the first way is doomed. The second way is the right tactic, and in sections 5.3 and 5.4 I argue that the belief-desire model is not as central as Smith presumes.

We do explain action by citing facts independent of the agent's psychological state. Even Smith accepts that we often explain action by citing a fact. He writes:

> . . . if a desire and a belief pair explains an action then it follows that the things that explain the desire and the things that explain the belief explain the action too. Thus, given that we can explain Fred's turning on the light by citing, *inter alia*, his belief that he can illuminate the room by turning on the light, and given that the fact that he can illuminate the room by turning on the light explains his belief, it follows that we can explain Fred's turning on the light by citing, *inter alia*, that very fact: that is, we can say that Fred turns on the light because he can illuminate the room by doing so. (Smith 1998, 19-20)

The fact, for Smith, is a normative reason to act. It is one which justifies the action and we can explain the action in terms of what justifies it. So far, it seems that Smith and I are on the same page.

Smith, however, goes on to argue that the belief-desire explanation still occupies the central position in explanations of action. He thinks that an explanation in terms of the relevant fact presupposes that there is a belief-desire explanation, which we can grant for now.[11] An explanation in terms of a belief-desire combination, however, does not presuppose the existence of a fact. If the agent is wrong about what

[10] Derek Parfit (1997) also defends a distinction between normative reasons as facts and motivating reasons as constituted by psychological states.

[11] As will become clear in Chapter 7, however, it is possible that an explanation in terms of a fact does not have an accompanying belief-desire explanation, such as if the action was done on the basis of an emotional experience. But this is jumping the gun.

the facts are, then she does not have a motivating reason of the kind I have endorsed, a fact. As a consequence, it would seem that we cannot give a true explanation of her action purely in terms of what she took to be a reason because the explanation would be false. But we can still give an explanation from her point of view in terms of her beliefs and desires.

For example, Jo goes to the library because her book is there. If it is the case that the book is at the library, then we could potentially explain her action by saying that she went to the library because her book was there. If it is not the case that the book is at the library, then Jo does not act in light of a fact. Yet, we can explain her action by saying that Jo believed that her book was at the library or that she desired to obtain the book which she believed was at the library. It is true that Jo believes that the book is at the library or that she desires to obtain the book, regardless of whether the library really does have the book. We can give a true explanation in terms of her states of belief and desire.

Smith takes the fact that all rational action can be explained in terms of states of belief and desire but not in terms of facts to show that belief-desire explanations are the central kind of rationalising explanation. He also takes this to show that motivating reasons are constituted by our psychological states. It would seem that an explanation of Jo's going to the library when her book is not there but she believed that it was, from her point of view, shows how the agent was acting for reasons even if she was not acting in light of facts. Therefore, her motivating reason is not a fact but is her psychological state. We might agree that something is lacking when the agent is wrong about the facts and, on Smith's kind of account, what is lacking is not a motivating reason but a normative reason. If Smith is right, then explanations in terms of facts are not the central kind of explanation of rational action.

We should resist this conclusion. It is necessary to emphasise what should be apparent by now, that the rationalising force of even belief-desire explanations lies in the way we explain by reference to how the agent takes herself to be acting for considerations that are independent of her psychological states. To quickly summarise some important points:

1. As we have seen in my discussion of Davidson, the agent's motivating reason for acting will not be her psychological states but what makes up the content of her psychological states.

2. We also saw that, if the agent has a reason, she needs to see what she does as desirable in some way. With Davidson, we saw that a primary reason gives the agent's reason for what she did because, amongst other things, it identifies

what the agent finds desirable. The rationalising force of such an explanation lies in the way the agent takes herself to have a reason.

3. Even someone like Smith accepts that what explains a belief or desire can also explain an action.

All of these factors together emphasise the central role that the agent's motivating reason that is not a belief-desire compound plays in a rationalising explanation. We should therefore try to find a model of explanation which gives the motivating reason its central place.

One way we can resist Smith's conclusion about the centrality of belief-desire explanations for rationalising action is by denying that we cannot explain all action in terms of facts. If all action can be explained in terms of facts, then it does not follow that the central kind of rationalising explanation is in terms of beliefs and desires. We thereby undermine the claim that motivating reasons are constituted by psychological states. This is the tactic that Jonathan Dancy takes and it is distinctly unsatisfactory.

Dancy maintains that the reasons which explain the action will be motivating reasons, as facts, but denies that reason explanations in terms of motivating reasons are factive (Dancy 2000, 131-37).[12] His test for factivity is to see whether we end up with a contradiction if we deny that things are as the agent takes them to be (132). If we do end up with a contradiction, then the explanation is factive. If we do not, then it is not. So, for Dancy, we can give the explanation:

(D1) Jo's reason for going to the library was that her book was there. But it wasn't.

To Dancy's ears, there is no contradiction. This, he argues, shows that the reason explanation is not factive. So, explanations in terms of the agent's motivating reasons will not be factive even though they may be true.

In cases of not acting in light of a fact, we can still explain the action in terms of the motivating reason. If the reason explanation in terms of the motivating reason is not factive, then it does not matter if the motivating reason is false as we cannot infer from the explanation alone that the motivating reason holds. However, Dancy does think that a better way to explain is to add additional contextual markers such as 'as she supposed'. We now get:

(D2) Jo's reason for going to the library was that, as she supposed, her book was there. But it wasn't.

[12] If a reason is factive, then we can infer that both p and q are true from 'the reason why it is the case that p is that q'.

Again, adding the 'But it wasn't' does not create a contradiction, for Dancy. This full explanation, however, is now factive because we do get a contradiction if we deny the full explanans:[13]

(D3) Jo's reason for going to the library was that, as she supposed, her book was there. But she did not suppose that her book was there.

If Dancy is right, then we can reject Smith's claim that belief-desire explanations are central because we can give true non-factive explanations in terms of the agent's motivating reasons even when the agent is wrong about the facts.

This proposal is intuitively uncomfortable. For one thing, we could have different contradiction-seeking ears to Dancy and reject that there is no contradiction involved in (D1) and (D2). Both Jennifer Hornsby (2007) and Clayton Littlejohn (2012) take this tactic, providing linguistic data to support that there is a contradiction. I share their ears but, as Littlejohn notes, the linguistic data could be overridden by powerful arguments. A deeper worry is that a non-factive explanation is insufficient.

One way it is insufficient is by passing on explanatory power to the contextual markers. According to Dancy, an explanation like 'Jo went to the library because her book was there' is not factive. It follows that even in the case where she is right and the book is at the library, we cannot infer this from the explanation. But we are now left with a very puzzling explanation: Jo went to the library because her book might or might not have been there and we cannot infer from the explanation which was the case. If we have to introduce further bits of information in order to infer what was the case, information such as 'as she supposed', then we get a better understanding of what happens with an explanation in terms of the agent's beliefs and desires because we actually come to understand what the case is. The contextual feature, in the factive version of the explanation, is doing the explanatory and informative work. It seems that Smith would be right about the centrality of explanations in terms of psychological states.

A second way it is insufficient is that it is a puzzle why we often revise reports of why we acted when we find out that we were wrong about the facts (Littlejohn 2012, 152). For example, if we ask Jo why she went to the library, she may say that the book she needs is there. On discovering that the book was actually checked out by someone else last week, she will typically revise her report to something like 'I thought that the book was there'. But if her original report was true, as Dancy

[13]The explanans of an explanation is that which does the explaining; the explanandum is that which needs to be explained.

claims, then there is no need to revise her report. Dancy's account cannot explain why this would be the case and entails a kind of error theory in the way we revise our reports when we do not need to.

Similar worries can be levelled at views which endorse that reasons are propositions which, if true, would favour acting, such as endorsed by Mark Schroeder (2008): if true, Jo's reason that the book is at the library would count in favour of going there. But this too will not work. First, we would not be able to construct explanations like (D1) without contradiction. It is still the case that Jo's reason is that her book was at the library, and if (D1) is contradictory when dealing with Dancy, it is contradictory when dealing with the propositional view. Second, similarly to Dancy's model, we would not be able to infer from an explanation like 'Jo went to the library because her book was there' what was the case. If we introduce markers like 'as she believed', then we can infer that Jo had beliefs. We get a better understanding of what happens with an explanation in terms of the agent's beliefs and desires. Finally, if the agent's reason is a consideration which, if true, would count in favour of acting, we are left puzzled by why the agent will revise her report when she finds out that she was wrong.[14]

Dancy's proposal and the propositional views do not show that Smith is wrong about the centrality of the belief-desire model. We must turn to the second way of challenging Smith's account, which is by denying that all action can be explained in terms of beliefs and desires.

5.3 THE LIMIT OF BELIEF-DESIRE EXPLANATIONS

In this section, I argue that not all action is the result of states of belief and desire, and so we cannot explain all action in terms of states of belief and desire. In particular, pure expressive action of the kind that makes up my Category C of examples of emotion explanations cannot be explained in terms of states of belief and desire.

As Rosalind Hursthouse concisely defines it, pure expressive action is intentional action that the agent did not do for a reason 'in the sense that there is a true description of action of the form "X did it (in order to)…" or "X was trying to…"'; and the agent would not have performed the action if she were not in the grip of an emotion (Hursthouse 1991, 59). It is action for which one might say, 'I ϕ-ed because I was so frightened (or happy, excited, ashamed… so overwhelmed by hatred or affection or…) that I just wanted to, or felt I had to' (58). It is not action done to communicate an emotion or to release tension. It is generally accepted that it

[14] I draw on Littlejohn's (2012) helpful comparison of views like these with Dancy's view.

is intentional, that is, that it is under the agent's awareness and control, to some degree.

Two examples of expressive action which I introduced in Chapter 2 were:

- In a wave of hatred, Jane tears at Joan's eyes in a photograph. Jane is not trying to communicate her anger nor relieve her pent-up energy. She acts purely because she is angry.

- Having lost his much-loved wife, the man rolls around and buries his face in her clothes. Like Jane, he is not communicating his grief or relieving his emotion; he is simply expressing it.

If all action can be explained in terms of states of belief and desire, including intentional action like this, the challenge is to show how. An initial attempt to explain expressive action in terms of states of belief and desire might seem to ascribe some rather bizarre beliefs and desires to an agent (Hursthouse 1991, 60). For example, Jane tears at Joan's eyes in a photograph out of anger. Because Jane is angry with Joan and not Joan's photograph, it looks like the explanatory desire will have to be something like: Jane desires to scratch at Joan's face. But then the associated belief would be something bizarre like the belief that scratching the photograph is scratching Joan. Such a belief is false, and it is also surely false that Jane has this belief in the first place. I shall therefore focus on more sophisticated accounts of explanations of expressive action, such as those of Smith (1998) and Peter Goldie (2000).

Smith proposes that, rather than ascribing a bizarre belief to the agent, we should look amongst the beliefs she actually has. What rationalises and explains the action of the man who rolls around in his dead wife's clothes, for example, is the desire to roll around in the clothes and the belief that he can. The role of the man's grief is to provide an explanatory context in which having those beliefs and desires makes some sort of sense. Smith writes:

> When we say that the man is rolling around in his dead wife's clothes because he is grieving for her we thereby locate the belief/desire pair that explains his action in a context in which having them makes a certain sort of sense. This is because grief at the loss of a loved one is, by definition, a state in which we are disposed to think, and to desire, and to do, all sorts of things . . . For it in effect tells us that the man is acting on the basis of some desire that it is typical for people who are grieving to have: in this case, the desire to roll around in his dead wife's clothes. (Smith 1998, 22-23)

I shall raise two general problems for this approach. In section 5.3.1, I focus on the first problem which is that the beliefs and desires that Smith identifies mischaracterise the action and trivialise acting for a reason. In section 5.3.2, I turn to the second problem which is that, even if the belief-desire pattern of explanation could be made to work, we would be left with a trivial explanation.

5.3.1 First problem: mischaracterising & trivialising the action

The first problem is that the states of belief and desire that Smith identifies mischaracterise the action and trivialise acting for a reason. I first focus on the desire aspect and show how it mischaracterises the action. I pose an objection to Smith, introduce Goldie's attempt to get around it, and show how that also fails. Turning to the belief aspect, I shall end this subsection by arguing that the kind of explanation we are left with trivialises acting for a reason.

As expressive actions are characterised, the agent would not have performed the action if she were not in the grip of the emotion. When we seek to explain her action, we want to explain why, given the emotional state she is in, the agent acts as she does. The problem with the desire aspect of Smith's proposal is that, if the explanation is partly that the agent desires to so act, we have provided hardly any illumination whatsoever.

Remember Anscombe's example of the dull babbling loon who wants the saucer of mud despite not seeing anything desirable in it. If the man scoops mud into a saucer, we can give an explanation of his action in terms of his states of belief and desire: he desires a saucer of mud, and believes that scooping mud into a saucer is a way of getting a saucer of mud. But let us compare the dull babbling loon to the grieving man. If the man's relevant desire is just that he desires to roll around in the clothes, and that is that, then his action is no different to the loon's. But surely there is a difference between expressive action of the man's and the nonsensical action of the loon's? There is, and we can turn again to Anscombe for insight into what it is.

Anscombe argues that to make sense of desire, we need to be able to characterise the object of desire as desirable. We cannot make sense of the loon case if the saucer of mud is not seen as desirable but the man still desires it. So, if the man rolling around in his wife's clothes acts because he desires it, we are left asking, 'But why did you want to do that anyway?' Unlike the babbling loon's answer to the mud case of 'I just wanted to', the grieving man does have an informative answer: 'Because I grieve.' Rolling around in the clothes is seen as desirable because, we learn, the man is experiencing an overwhelming and very sad emotion and emotions often

call for expression. Grief therefore provides a further characterisation of the action, making the action intelligible because it shows why the agent saw it as something to do.

In addition, our concept of grief is of a deeply felt and motivating state. So, not only does the grief make the action intelligible, it also shows how the agent saw the action as desirable because it is an expression of a deeply felt and motivating state.[15] The grieving man is different to the dull babbling loon precisely because there is this further characterisation to why he finds the action desirable and to say that the man simply desires to roll around in the clothes mischaracterises the kind of pro-attitude that he has, as one without a further desirability characterisation which it does in fact have. We thus see that the man's grief does more than simply set up a context where certain beliefs and desires make sense; it is fundamentally part of why the agent sees the action as desirable in the first place. The man's action is quite different to the loon's but the belief-desire model overlooks that difference.

Goldie also finds Smith's proposal unsatisfactory because of a similar complaint. Returning to the photograph-scratching example, Goldie writes, '[i]t does not really make it at all clear why someone should be disposed when hating or in anger to desire to do such a "bizarre" thing as scratch the eyes in a photo of the person they hate or are angry with' (Goldie 2000, 128). He allows that some desires are 'primitively intelligible' if 'they cannot be better explained in virtue of anything else'. An example of a primitively intelligible desire would be: desiring to get revenge on someone with whom you are angry. The problem with expressive action like in the two examples under consideration, for Goldie, is that the desire to scratch out the eyes in a photograph or the desire to roll around in someone's clothes is not primitively intelligible, even in the context of an emotion.

Goldie proposes that we should rather understand the action as an expression of a wish. The expression involves being disposed to imagine some desire being satisfied. So, working with the example of Jane scratching out the eyes of Joan in a photograph, the explanation is: 'Jane has a desire to scratch out Joan's eyes, and she imagines that she is doing this in the expressive action' (Goldie 2000, 130).

For Goldie, the action is explained along two 'vectors'. The action vector is explained in terms of the belief-desire pair that Smith identifies, and the expressive vector is explained with reference to a desire and a wish. We make sense of the bizarre desire to scratch out the eyes in a photograph, contained in the action vector, by appealing to a more intelligible desire to scratch out the actual eyes, accompanied

[15] Even if the grief provides a desirability characterisation for the action in this way, it does not do so by showing how the action was desirable *in order* to achieve anything.

by the wish or the imagining that this is what one is doing. This desire is in turn primitively intelligible in the context of the emotion of anger. Jane is angry with Joan and scratches out the eyes in a photograph of Joan because it is a photograph of Joan, and she wishes to scratch out Joan's eyes. Similarly, the man is grieving and rolls around in his dead wife's clothes because they are her clothes and he wishes to be with her.

As Goldie notes, his account does not apply to all cases of expressive action. The cases where it does not apply are when there is no clear relation between the object of the emotion and the object of the expression, like there is between a person who is the object of one's anger or grief and their photograph or belongings. In such cases, the explanation of the action is simply that the person is in the grip of the emotion. For example, while arguing with his wife, a man smashes a vase which belongs to her 'simply because he is in the grip of anger; there is no further explanation beyond the primitively intelligible desire to vent one's anger' (Goldie 2000, 134-45).

The fact that Goldie's account does not apply to all kinds of expressive action, however, highlights that an expression of a wish is not a truly expressive action of the type we are seeking to explain.

If you have a wish you want to enact, then there is a right and a wrong way, or a successful and unsuccessful way, to enact it. But, as Hursthouse points out, pure expressive actions do not have the possibility of being expressed in the wrong way (Hursthouse 1991, 60-61). An expression may be inappropriate or disproportionate but it is not wrong as an expression. She levels this as a fault with the belief-desire model: if you have a belief about what description your action falls under, then there must be the possibility that your action is the wrong way to do things. My objection is similar: if you have a wish, then there is the possibility that your action is the wrong way to enact it.

To illustrate, consider the man who smashes his wife's vase and suppose that it transpires that he actually smashes his brother's vase. In the purely expressive case, this makes no difference to the 'success' of his action as an expression of his anger. He acts expressively and that is that. In the expression of a wish case, in contrast, it does make a difference. The symbolic connection between his wife and her vase is no longer there. His enactment of a wish is thwarted.

Similar scenarios can be constructed for Jane and the grieving man. On discovering that the photograph is not of Joan, Jane's expression is not affected in the purely expressive case: it is sufficient that the photograph looks like Joan but Jane could also express her anger by gouging holes in her notebook. In the wish case, her action has backfired: she has not achieved what she wanted, which was to enact

scratching out Joan's eyes on an image of Joan. The grieving man still acts expressively even if the clothes turn out to be his daughter's; he may roll around in the clothes because they smelt like his wife, they looked like her clothes, but that is not why he expresses his grief. The man expressing a wish, however, has failed in the enactment because the clothes are not his wife's and rolling around in them does not enact his wish to be close to her, even if the clothes look and smell like her clothes.

What about cases where there is no symbolic connection but the agent is still expressing a wish? For example, the man angry with his wife wishes to harm her and grabs any substitute and tries to break it. He is enacting the wish of harming his wife. But suppose that the object does not break. He has failed in his symbolic act of harming his wife. In contrast, if the man were purely expressing his anger, not breaking or harming the object does not affect the 'success' of his expression. He may feel foolish that the object does not break, but from this it does not follow that his expression was not successful. Rather, he may realise the extremity of his expression and feel foolish about feeling so strongly that he acts as he does, or, perhaps feel foolish about the way in which he is just so angry he cannot even successfully break a vase.

These examples show that the agent is acting in order to achieve something when she acts from a wish. The explanation is of the form 'X ϕ-ed in order to symbolically enact a wish'. Therefore, it is not the kind of expressive action we are seeking to explain.[16]

While Goldie does describe a set of interesting cases, his account focuses on the wrong ones and is thus inadequate. However, even if it were addressing the right set of cases, the account faces other problems. This is because it relies on the belief-desire explanation of action that Smith proposes, and so it faces the same problems as Smith's does, to which I shall now return.

Turning to the belief aspect, we again run into problems with the characterisation of what the man takes himself to be doing when he rolls around in his dead wife's clothes.

In order to have a primary reason, which is what rationalises the action for Smith, the agent must have a pro-attitude towards a class of action, such as rolling

[16]Sabine Döring (2003) poses a similar objection. She argues that Goldie's account faces a dilemma where either the explanation of expressive action is available in terms of a belief-desire account or it is not. If the former, then where there is symbolic value attached to the object of the action, the agent is not acting purely expressively. (Döring does not address the wish-aspect of Goldie's account, but we can bolster this horn of the dilemma by including that the agent is enacting a wish.) If the latter, then Goldie has not shown what he set out to do, which was to give an account in belief-desire terms.

around in the clothes,[17] and he must have the belief that this action, under a certain description, falls under that class. If so, then the agent's reason will be composed of the belief along the lines of, 'rolling around in my wife's clothes is a way in which I can roll around in my wife's clothes'. The reason explanation, then, is that the man rolled around in his wife's clothes because he wanted to and believed that doing so was a means of doing so. This is distinctly unsatisfying as an explanation and trivialises the notion of acting for a reason.[18]

5.3.2 Second problem: trivial explanations

The second problem is that, even if the belief-desire pattern of explanation could be made to work, we would be left with a trivial explanation.

We see this by the fact that different emotions can set up different contexts for the same beliefs and desires. For example, another man is rolling around in his dead wife's clothes. Rather than grieving, he is ecstatic that his evil and domineering wife has died and that he is finally free. He rolls around in her clothes because he desires to do so and believes that he can. On a belief-desire model, the grieving man and the ecstatic man act for the same reason and we explain their action with the same reason. If that is where the explanation leaves us, then it is trivial.

Smith is right that the two men's emotions provide a context for the action. But what is really of interest in explaining the action is identifying what it is about *grief* and what it is about *ecstasy* that do so. Unlike the man who wants a saucer of mud, the two men who want to roll around in the clothes see the action as desirable in some way.

In these particular cases where we have one man rolling around out of grief and

[17] One might object that the relevant pro-attitude is towards a class of actions which will get the man closer to his wife. But this cannot be the case if we are dealing with purely expressive action. If that is the relevant pro-attitude, then the man is acting in order to achieve something, namely to get closer to his wife, and the action is no longer purely expressive.

[18] Alvarez (2009) makes the same point with a non-emotion example of doing a cartwheel, and with regards to saying that expressive action is acting just because I want to. She writes:

> If we apply this [Davidsonian] model to our example, we get: "I did a cartwheel because I had a pro-attitude (e.g. a desire) towards actions with a certain property (namely, the property of 'being a cartwheel'), and also had a belief that my action A, under description d (namely, 'doing a cartwheel'), had that property"; or more conversationally, "I did a cartwheel because I wanted to do a cartwheel and believed that my action of doing a cartwheel was *what?* a means? an instance? a way? of doing a cartwheel". But this barely makes sense, and the only reason to claim that I had such a belief is that then the example will fill the Davidsonian model. (Alvarez 2009, 299)

another rolling around out of ecstasy, the difference is in what the emotions are about and in how the actions are seen as desirable. Both emotions are about the dead wives. But the man who is grieving sees the loss of his wife as irrevocably bad, whereas the ecstatic man sees the loss of his wife as a release from tyranny. The first is drawn to the action as an expression of grief, the second as an expression of ecstasy. If Smith is to allow that the emotion provides a context, and he is right to do so, and if the kind of context an emotion provides is dependent on how the agent experiences things emotionally, then the emotion provides a context in virtue of its intentional quality and intentional nature. But, as I argued in Chapter 4, emotions are not reducible to belief-desire compounds. So, why should we side-line the emotion in the way that Smith does, especially if the explanatory focus is shifted back to the emotion? Without giving an emotion its central place, a belief-desire model of explanation is trivial. But if we give an emotion its central place, it is the emotion as an intentional state which explains, not some belief or desire. Smith's account does not provide a satisfactory explanation of expressive action.

Goldie attempted to account for the role of the emotion by arguing that emotions involve primitive desires. The action is still explained in belief-desire terms, but the expressive nature of the action is explained with reference to a desire made primitive by the emotion, together with a wish to satisfy that desire. The wish is symbolically enacted. As I argued, the class of action that is expressive of a wish is still action for a reason and so he has not provided an account of the nature of our target actions. Even if it did apply, Goldie's account falls into the same quagmire as Smith's with regards to trivialising both acting for a reason and giving a trivial explanation. We must conclude that expressive action is not action done for reasons that are combinations of beliefs and desires. Consequently, we cannot explain such action as rational in terms of beliefs and desires. It therefore seems not all actions are done for beliefs and desires and we need another way of explaining expressive action.

Several philosophers have argued that expressive action is not action done for a reason or a reason of the 'in order to' kind.[19] Hursthouse led the way with her defining paper 'Arational Actions' (1991) where she argues that expressive action is intentional action that is arational. Arational action is action not done for a reason but which is not irrational. It is simply arational. While it is arational, she thinks that it can be given a rational or irrational description, and it is this description which makes us think of some expressive actions as either rational or irrational, not the

[19]For example, see Hursthouse (1991); Döring (2003); Alvarez (2009); Betzler (2009).

nature of the action itself.

The failure of attempts to show how acting purely expressively is acting for a reason, the very nature of such actions—that they are somehow different to paradigmatic action for reasons—and the existence of convincing accounts of the difference between intentional and rational action is good support for supposing that such actions are not actions done for a reason at all. Nevertheless, they are intentional and not irrational. They are arational. I find these considerations compelling and shall accept that expressive action is arational.

This means, however, that the way in which we explain expressive action need not have any direct bearing on how we explain rational action. If we want to explain action done for a reason and if we cannot explain action not done for a reason in terms of states of belief and desire, it does not follow that action done for a reason is not rationalised by states of belief and desire.

What does follow is that states of belief and desire are not required for an action to be intentional. I ended the previous section by drawing attention to Smith's claim that the belief-desire model of explanation is central *because* all intentional action can be explained according to it. But not all intentional action can be explained according to it, and so we do not need to accept that the belief-desire model of explanation is central. We also saw that a belief-desire explanation risks trivialising the notion of acting for a reason if there is not some further desirability characterisation of the action. What makes the action desirable, however, is not a state of belief or a state of desire, but some consideration or fact—a motivating reason.

This means that, even if we can give an explanation of action in terms of states of belief and desire, it may still be the case that they are not necessary for an explanation and we can give another form of explanation of the same action, one which gives the motivating reasons as facts their central place in rationalising action.

5.4 Motivating reasons and explanatory reasons

In this section, I propose and adopt an alternative model for explanations of rational action in terms of the agent's motivating reasons as facts, one which avoids the problems facing Dancy's proposal. Alvarez (2010) distinguishes between motivating reasons and explanatory reasons and, by making this distinction, we do not have to commit to the idea that the reasons the agent takes herself to have and the reasons that explain her rational action are always the same, and neither do we have to give up the factive nature of explanations.

Explanatory reasons are more general than motivating reasons. An explanatory

reason 'plays the role of (part of) the explanans in an explanation of the agent's ϕ-ing' (Alvarez 2010, 35). While an explanatory reason can still provide an explanation from the point of view of the agent and why she did what she did, an explanatory reason need not pick out the actual reason for which the agent took herself to be acting. We might explain with reference to an explanatory reason when the agent is wrong about her reasons or when we wish to remain neutral on whether her reasons were real, such as when Jo goes to the library because she believes that her book is there but it is not. We may also explain by giving other features of the agent which give us insight into her point of view, such as a character trait. For example, if James does not go to the party because he is shy, his reason for not going to the party is not that he is shy even though we can explain his action in terms of his shyness.

A motivating reason will always be an explanatory reason because it provides an explanation in virtue of being the reason the agent took herself to have. The reverse need not always hold. If the explanatory reason is that Jo believed that the library had her book, and if she is wrong, we cannot explain simply by saying that the library had her book.

Nevertheless, we can still explain her action from her point of view by identifying an explanatory reason which picks out what the agent took to be a reason. Because we are explaining Jo's action from her point of view, we are still explaining her action as a reasonable or potentially as rational action. This allows us to distinguish between psychological explanations in terms of beliefs and desires which pick out what the agent took to be a reason, and psychological explanations which explain in other ways, such as in terms of a character trait like being shy. A belief-desire explanation still gives an explanation in terms of what the agent's considerations in favour of acting were, even if she was wrong about those considerations being reasons. On this kind of framework, we could understand Smith's 'motivating reasons' as explanatory reasons.

One may still object that a motivating reason is not actually what rationalises the action. The idea behind this new form of the objection is that, in cases where the agent is wrong about her reasons, there are still considerations in favour of why the agent did what she did: Jo did not just randomly find herself at the library, she went there with a purpose and in order to achieve a goal. She is completely reasonable in acting as she did, given the beliefs and desires she had, and what rationalises her action is showing that she took herself to be acting for a reason, not that she actually did act for a reason.

On this line of thought, we might be tempted to distinguish between 'subjective'

and 'objective' reasons. Subjective reasons are 'what the agent *justifiably sees* as a reason because of the contents of his actual mental state' (Döring 2010, 287). Objective reasons are what are reasons because of features of the agent's situation. The claim, then, is that what rationalises an action is that the agent acts for what she believes is a reason, regardless of whether the reason is a fact.

Sabine Döring (2007, 287) gives the example of being told by a doctor to take a certain medicine. Unknown to you, the medicine will actually harm you. In such a case, the objective reason is that the medicine will harm you and so you should not take it. If you believe that you ought to take the medicine because the doctor told you to and yet do not take it, you are acting irrationally even when what you have most objective reason to do is not take the medicine, Döring claims. This suggests that what actually makes an action rational is that you are acting in line with your beliefs about what your reasons are, not that you are acting in line with what the actual reasons are.

If this is the case, then an explanation in terms of the agent's beliefs and desires may be sufficient to rationalise an action and, even stronger, giving an explanation in terms of the objective reason alone does not rationalise the action because it does not show how the agent acted for what she believed was a reason.

We can reject this kind of argument for three reasons.

First, the objective-subjective reason distinction is not needed if we make the motivating-explanatory reason distinction. Giving an explanation in terms of the agent's motivating reason is not the same as giving an explanation in terms of an objective reason, even though a motivating reason is an objective reason. This is because a motivating reason is the reason in light of which an agent acted, not just any good-making feature of the situation. Even if it is the case that the medicine will harm me, and it is a reason for which I ought to act, if I am not aware that it will harm me, it cannot be my motivating reason. This means that the stronger claim must be rejected: if we explain an action in terms of the motivating reason, it is implicit that the agent was relevantly aware of the motivating reason, perhaps by believing that it is the case. It is therefore compatible with rationalising an action by identifying the motivating reason that we are showing that the agent is acting for what she believes she has reason to do.

Acting for a motivating reason does not imply that one is acting for the best reason or, even, a good reason. One can act guided by a fact which is a good reason for the action, but one can also act guided by a fact which is a bad reason. For example, I go to the corner shop to buy milk because it is a shop which stocks milk. This is a good reason to go to the shop to buy milk. But perhaps I go to the corner

shop to buy milk because I dreamt last night that this is what I would do today. This is a bad reason to go to the shop to buy milk. I can also act guided by a false consideration which, if true, would be a good reason or a false consideration which, if true, would be a bad reason. The shop is out of milk. Nevertheless, if it had milk, that would be a good reason for going there to buy milk. I did not dream about buying milk from the corner shop. Even if I had, it would be a bad reason for going to the shop to buy milk.

Second, the subjective-objective reason distinction cannot account for the normative force of reasons to act. Imagine that the building you are in is burning down. If there is a *pro tanto* reason for you to ϕ, then you ought to ϕ; and there is a *pro tanto* reason for you to get outside, namely that the building is burning down. So, you ought to get outside. But if having a subjective reason is required for acting rationally then, if you do not believe that you have a reason to get outside, then you do not have a reason to get outside on pain of irrationality. But there is a reason, which is the fact that the building is burning down. How do we adjudicate between what you should do in this scenario? We cannot say that the normative force of the *pro tanto* reason to get outside is defeated by your lack of belief that it is a reason regardless of how well that belief coheres with your other beliefs, else we lose the normative force of reasons. At best it seems that we can say there is a reason for you to believe that the building is on fire, namely the fact that the building is on fire, but that same fact is not a reason for you to act. This is an undesirable conclusion to be forced into, because we do think that reasons have normative force on our actions and our acting rationally, independent of whether we believe them or not.

Third, we can accept that not taking the medicine despite believing that you should is problematic. However, the problem need not arise from your not acting in lines with your beliefs. What is problematic, even irrational, is that your beliefs and attitudes are in conflict because you cannot, without contradiction, hold all of them together. Your belief that you should take the medicine is in direct conflict with your intention not to take it; and in cases of such conflict, you are rationally required to give one of them up.[20] This could be the belief or the intention but in the case as described, it seems that neither is given up. Hence, we have an alternative

[20]This is a wide-scope interpretation of rational requirements, introduced by John Broome. See Broome (1999, 2005) but also Kolodny (2005) and Way (2009) for discussion. On such an interpretation, rational requirements govern combinations of attitudes, not individual attitudes. In contrast, a narrow-scope interpretation would have that rational requirements govern individual attitudes. Here is the contrast. On wide scope, if I believe that I should take the medicine, then I must either take the medicine or give up the belief. On narrow scope, if I believe that I should take the medicine, then I must take the medicine. The wide-scope interpretation is widely accepted (Wallace 2014).

explanation for the seeming irrationality, one that is independent of the subjective-objective reason distinction.

Nevertheless, the subjective-objective reason distinction does capture an important intuition, but we can capture that intuition by adopting alternative distinctions which are not similarly problematic. I shall distinguish between acting rightly or in light of a reason and acting reasonably, and subjective and objective rationality.

'Acting rationally' is an ambiguous term. As I have been using it, acting rationally is acting in light of a reason. A reason is a fact that counts in favour of the action so, if one acts rationally, one is doing what one has reason to do. In this sense, acting rationally is acting rightly. One can also use 'acting rationally' to refer to acting reasonably. Acting reasonably need not be acting rightly or in light of a reason, but it is acting in a way that is warranted, considering the beliefs and desires that you have.

To see the distinction, consider an example that Littlejohn (2012, 149) gives to show that one can act reasonably but not rightly (in light of a reason). Parents have to decide when it is all right to allow their children to play on their own in the neighbourhood. When it is all right will depend on what the neighbourhood is: in safe neighbourhoods, the parents may end up being overly protective but in dangerous neighbourhoods, insufficiently protective. Given the variety of factors, parents may reasonably but incorrectly judge when they ought to let their children out. If we insist that acting rightly (for a reason) and acting reasonably are the same, then we would be forced to say that parents cannot reasonably but incorrectly judge when to let their children out, but that is surely false.

We now see that the subjective-objective reason distinction is not without basis. What the distinction attempts to capture is the idea that acting reasonably may not be acting in light of a reason. However, as we have seen, there are problems with making the distinction one of types of reasons. We should rather adopt Niko Kolodny's distinction between 'subjective' and 'objective' rationality as a means to talk about the intuitive difference (see Kolodny 2005). Subjective rationality refers to the relations that hold between our psychological states and attitudes, whereas objective rationality refers to the relations that hold between our psychological states and facts. Rather than making a claim about different kinds of reasons and thereby running into problems regarding the normative force of having a reason, we can talk about different kinds of relations between states and facts. Acting reasonably is acting coherently with our other attitudes: it is being subjectively rational. In good cases, we also act rightly and in line with the facts: we are objectively rational.

The subjective-objective rationality distinction also helps us to explain what is

intuitively compelling about a psychological explanation in terms of beliefs and desires. What is intuitively compelling is not that it shows that the agent acted for a reason. Rather, it shows that the agent was subjectively rational and acting reasonably. Compare this to an explanation in terms of a character trait like shyness. Here, the explanation does not show that the agent was subjectively rational, that she acted reasonably or that she acted for a reason.

We therefore see that by distinguishing between motivating reasons and explanatory reasons and by allowing that they can sometimes overlap but need not always do so, we both keep motivating reasons central in a rationalising explanation and avoid the problems of Dancy's proposal. Dancy insisted that the reasons which explain will always be the motivating reasons, but this is not the case. We can therefore give explanations in terms of beliefs and desires and give explanations in terms of motivating reasons but, because explanations in terms of beliefs and desires explain by showing how the agent took herself to have a motivating reason, the latter kind of explanation is the central kind. Finally, explanations in terms of the agent's beliefs and desires are not completely disconnected from the agent's rational action. Such explanations can still show how the agent was subjectively rational and acting reasonably.

5.5 Conclusion

I set up the discussion of this chapter by asking what a rational action is. I defined it as action that is under the agent's awareness, control, and is action that is done in light of a reason. I then argued that acting for a reason is acting for some fact in light of which the agent sees that action as good. That fact is stated by a true proposition and is the agent's motivating reason.

Turning to explanations, I argued that, while we can explain a rational action in terms of states of belief and desire, those states of belief and desire are not the motivating reasons. I showed how we can explain an agent's action by identifying the motivating reason and argued that such an explanation rationalises an action. If the agent was wrong about her reasons then, while we cannot explain the action as one done for a reason, we can still show how it was reasonable by showing how the agent took herself to be acting for a reason.

The discussion in this chapter helps to formally describe what the differences are between the kinds of explanations in the different categories I introduced in Chapter 2, and to phrase more precisely what is so different about Category E. These categories of explanation were:

A	Manifestation	A1. Symptoms of manifestation A2. Manifestation in behaviour
B	Other intentional states	B1. Beliefs and desires B2. 'Seeing' B3. Motives
C	Pure expression	C. Expressing an emotion
D	As a reason: addressing a state	D1. Avoiding, alleviating, encouraging: (a) Symptoms of manifestation (b) Manifestation in behaviour D2. Avoiding, alleviating, encouraging the emotion, not just the manifestation
E	Providing a reason: intentional nature	E. Not avoiding, etc., or expressing

Category A illustrates cases where the emotion plays a causal role but is not the reason in light of which the agent acted. We can now more formally say that the kind of explanation in Category A is not a rationalising explanation.

The explanations in Category B are either examples of cases where the emotion leads to or is associated with other beliefs, desires and attitudes of the agent. The kind of explanation thus shows how the agent had the beliefs, etc., she had. Or, as in the case of the explanation referring to the agent's motive, the explanation draws on how emotions have associated mental states, etc. The kind of explanation is thus not a rationalising one but it is one which potentially shows the action as reasonable given the agent's other states and attitudes. The emotion in the explanation is an explanatory reason.

In this chapter, I have argued that the examples in Category C are examples of arational action and so the explanations, too, are not rationalising explanations. The emotion does not play a rationalising role and it is an explanatory reason in the explanation.

This leaves categories D and E as potentially illustrating how an emotion can play a rationalising role of some kind. I argued in Chapter 3 that the emotions in the examples in Category D are the agent's reasons because they are states to be addressed. We can now say more formally that they are the agent's motivating reasons: the emotional experience, as an experience, is a consideration in light of which the agent sees the action as good. These cases are not different from other rationalising explanations despite the involvement of an emotion.

From the discussion of this chapter, however, we can now see how the explana-

tions in Category E could be quite different in kind to the explanations in the other categories. There are two important ideas that can be used to crystallise a description of the kinds of cases in Category E. The first is that, if an agent's reason is not her states of belief and desire but a fact in light of which she sees the action as good, then another intentional state which alerts her to the fact could provide the reason. That other intentional state could be an emotion. For example, if Richard buys flowers for Katherine because he loves her, his love need not be his reason in a way like it would be in Category D—a state to be addressed—but the information which the emotion provides, namely that Katherine is affection-inspiring and valuable in his life, could be his motivating reason. The second idea is that we do not need states of belief and desire in order to explain an action, then other psychological states could also explain an action as reasonable, much like states of belief and desire can. If Richard bought flowers because he loved Katherine then, while the emotion in the explanation is an explanatory reason, the nature of the explanation is one which shows that he acted for a putative reason and acted reasonably. Establishing whether such an explanation is possible is the task of Part III, to which I now turn.

PART III

CHAPTER 6

THE INTENTIONAL NATURE OF EMOTIONS

In Part II, I argued in Chapter 4 that emotions are intentional states which provide us with information about how things are but which are neither belief- nor judgement-like. I also argued in Chapter 5 that acting in light of a reason is acting for a motivating reason, which is a fact, and that giving a rationalising explanation requires showing that the agent had a motivating reason. We have now canvassed the basis to articulate the view that, if we act on the basis of an emotional experience, then we can act in light of a reason, perhaps in a way much like how we act in light of a reason when we act on the basis of a belief. What remains now is to show how, in virtue of their intentional nature, emotions can provide access to reasons in light of which the agent acts and that when she acts on the basis of an emotional experience, she can act in light of those reasons. We need to establish what, exactly, those reasons are and whether we can indeed act in light of them in the right kind of way. This is the task of Part III.

In this chapter, I examine the concept of the formal object of an emotion more closely than I have so far done. This is because it is central to the idea that emotions are intentional and that emotions could provide access to information about how things are, information in light of which we act. In other words, it is central to the idea that emotions could provide access to reasons.

I begin in section 6.1 by examining the concept of the formal object and how it is used in the emotion literature. In section 6.2, I examine what characterisation of the formal object best meets the conditions and requirements we have of it and I argue that the concept of the formal object is best understood as a response-dependent concept. However, response-dependence brings with it the worry that the formal object may now be trivial because it becomes something like 'whatever property elicits emotion x in subject S'. I therefore argue, in section 6.3, that the formal object is not trivial if understood in this way because we can be wrong about when it

obtains and because we can fail to respond emotionally to it; that is, it is not the same as 'whatever property elicits emotion x in subject S'. In light of this understanding of the formal object as a response-dependent concept, in section 6.4 I articulate three potential reasons to which an emotion can provide access in virtue of its intentional nature. These are:

1. The fact of how things seem to the agent.
2. The fact of how things are, as in what response is merited.
3. The fact of how things are, as in what relation holds.

6.1 THE FORMAL OBJECT OF AN EMOTION

Part of our understanding of emotions as intentional states is that they are about particular objects at which they are targeted. But, as we saw in chapters 1 and 4, it is supposed to be in virtue of some formal object that the particular object is the object of the emotion it is. In this way, the concept of the formal object plays an important role in enhancing our understanding of emotions because it is the concept of something that illuminates the nature of each emotion.

The formal object of a mental state is something, generally taken to be a property, that is suitably general so that a range of particular objects can have it. We might say, for instance, that the formal object of belief is truth, where truth is something general which a proposition, the particular object of belief, can have. Further, a mental state is a belief in virtue of being a mental state which takes a proposition as true, rather than imagining a proposition as true, wishing it were true, etc. This entails two ways of assessing a belief. The first is to assess whether the proposition really is true, and is a question about the formal object as a property. I shall use 'formal object as a property' and 'formal property' interchangeably. The second is whether the mental state is what it is in virtue of taking the proposition as true, rather than one of imagining it as true, wishing it were true, etc., and is a question about how the formal property relates to the mental state. It is a question about the concept of the formal object. The actual truth of the proposition does not affect whether the mental state is a belief or not because we can believe things that are not true. We also do not have to believe everything that is true.

In a similar fashion, the formal object of an emotion must be a property that a range of particular objects can have. It is generally taken to be an axiological or

evaluative property.[1] So, the formal object of an emotion is an evaluative property which the particular object of an emotion can have. Further, an emotion is fear rather than hatred, for example, in virtue of being elicited by the property that is fear's formal object. And, like belief, we can assess an emotion by asking questions about both the formal property and the concept of the formal object.

For instance, we can ask two distinct questions about the formal property. Like for belief, we can ask whether the particular object really has the relevant property. But we can also ask whether something else which is *not* the particular object of the emotion has the property, in virtue of which the emotion is elicited.[2] For example, I may be crying over spilt milk but it is in fact my grandmother's death that I have been ignoring which elicits my sadness. I am wrong in what I attribute the formal property to, but nevertheless right that something is that property, namely a significant loss. In this way, I can learn from my emotions. Turning to the concept and how the formal property relates to the emotion, and like for belief, we can ask whether the emotion is what it is in virtue of being elicited by the formal property. It is in virtue of being elicited by my grandmother's death as a significant loss that my emotion is grief rather than anger, for example.

We thus have two sets of questions through which we can increase our understanding of the intentional nature of emotions, questions about the formal property and questions about the concept of the formal object, including about the way the formal property relates to the emotion. We can start examining the formal object by asking: (1) what is the formal property itself; and, (2) how does the formal property relate to the emotion, and what role does it play in the intentional nature of an emotion?

So, what is the formal property itself? There is a lack of consistency in the philosophical literature about what kind of property the formal object of an emotion is. Some talk of the formal object of fear, for instance, as danger (Lazarus 2001)[3], the dangerous (Prinz 2004), a threat (Roberts 1988), the property of being

[1] See Fabrice Teroni (2007, 396) for discussion on this point. It is worth noting that emotions like wonder or surprise do not seem to have formal objects that are evaluative properties. Whether or not all emotions have evaluative properties as formal objects would affect whether 'emotion' is a natural kind but does not affect the claim that all emotions have a formal object and, more-often-than-not, it is evaluative. See also Kevin Mulligan (2010) for a discussion of value and evaluative properties, and different proposals about their connection to emotions. Mulligan argues, and I agree, that the formal object of an emotion is not a value although it may still be an evaluative property.

[2] I talk of the formal object as the property in virtue of which an emotion is elicited. The kind of elicitation can remain vague. An emotion could be elicited by automatically responding to some formal object, imagining a scenario with the formal object, believing falsely some proposition about the formal object, etc.

[3] Strictly speaking, Richard Lazarus talks about the core relational themes and not the formal

an aversive possibility (Roberts 2003), the threatening (Nussbaum 2001). These are evaluative relational properties, as described in Chapter 1. Others, however, talk about the formal object of fear as the frightening (de Sousa 1987), or the fearsome (D'Arms 2005; Tappolet 2010). This latter way of talking about the formal object is often interpreted as saying that the formal object is a response-dependent property.[4] And sometimes we find both as, while talking of fear as construing a threat, Robert Roberts (1988) also refers to the fearsome.[5]

We will not immediately do any better in answering the second question, about how the formal property relates to the emotion and what role it plays in the intentional nature of an emotion, because there is also a lack of clarity and agreement here. In order to pinpoint where the lack of clarity arises, let us return to belief.

Just because the formal object of belief is truth, it does not follow that my belief is about the proposition 'x is y is true' rather than, simply, 'x is y' which is represented as being true. In other words, the formal object of belief does not explicitly enter into the intentional content of the belief. It is how the explicit content, 'x is y' is represented as being and we can express the content in the *description*: the truth that x is y. So, just because the formal object of fear, say, is the frightening, it does not follow that the content of fear is 'x is frightening' rather than 'x' which is represented as frightening by being perceived or experienced in a certain way. We can express the content in the description: the frightening x.

Following the failure of strong cognitive accounts of emotions, no-one these days (of whom I am aware) proposes that the formal object needs to be explicitly represented. This is because explicit representation would require application of a concept, which in turn requires language and mastery of abstract concepts like the dangerous or the frightening. Nevertheless, there are contrasting views about how the formal object relates to the intentional content and this is where the lack of clarity arises.

Some suggest that the formal object makes up part of the emotion's own intentional content. Jesse Prinz, for example, talks about emotions being about their core

objects. However, amongst appraisal theorists in psychology, like him, the core relational themes are roughly equivalent to the philosophical formal objects (Deonna and Teroni 2012, 51).

[4]I will distinguish between response-dependent properties and response-dependent concepts in section 6.2.2.

[5]Christine Tappolet (2010, 328, fn 7) gives a list of these different identifications, which I have added to, and flags that one can interpret the formal object as relational or as response-dependent. Mulligan (2010) also identifies different forms the formal object has taken, such as being a value, an axiological or evaluative property, or a deontological property. I do not think that anyone really identifies the formal object as a value and so shall not focus on that option. The deontological property corresponds to the response-dependent property.

relational themes or formal objects. For him, emotions represent the formal objects and the formal objects are represented in the intentional content (Prinz 2004, 62). For Christine Tappolet (2010), emotions have intentional content that is an evaluative but non-conceptual representation of the formal object. As she writes, '[t]he idea is often formulated as the claim that emotions are perceptions of values, where it is understood that the perceptions in question are non-conceptual' (Tappolet 2010, 329).[6]

Others suggest that the formal object need not be in the emotion's intentional content. On Richard Lazarus's brand of appraisal theory, for instance, the core relational themes—his equivalent notion to the formal object—are not explicitly represented. They are what emotions are designed to represent, but they are only implicitly represented by other explicit representations in the appraisal dimension, such as appraisals about goal relevance, goal congruence, ego-involvement, coping potential, etc. (Lazarus 1991, as discussed in Prinz 2004). This suggests reading the role of the formal object in the intentional nature as a description of the emotion's intentional *nature* as a whole, rather than a claim of what the emotion's actual intentional *content* is. Kevin Mulligan (2010) argues that emotions do not represent the formal object, even while the formal object is what justifies or explains an emotion. And Julien Deonna and Fabrice Teroni (2012) argue that the way in which emotions have correctness conditions traces back to the kind of attitude emotions are rather than their specific content. Indeed, they argue that emotions are not about evaluative properties even if the particular objects are bearers of evaluative properties. We therefore find differing views and a lack of clarity about the role of the formal object in an emotion's intentional nature.

In the next section, I examine the concept of the formal object in more detail in order to clarify what it is and what role it plays in the intentional nature of an emotion. I argue that the the concept of the formal object is a response-dependent concept because it is the concept of a property which produces experiences of a certain kind in a responder. Nevertheless, emotions need not be explicitly about the formal object under this full description, even though the emotion's intentional nature as a whole represents its formal object and we can describe the implicit content in terms of the formal object.

[6]While Tappolet says that emotions are perceptions of values, I interpret her as referring to evaluative properties.

6.2 THE REQUIREMENTS OF THE FORMAL OBJECT

In order for the concept of the formal object to enhance our understanding of emotion, there are two conditions which our characterisation of the formal object must meet. First, it must be possible to experience an emotion even when the relevant formal object is not the case or what elicits the emotion. Like for belief, whether the property actually is what elicits the emotion or not does not affect whether the emotion is what it is, so long as we experience the property as obtaining and as eliciting our emotion. We can and do experience inaccurate or inappropriate emotions. Second, it must be possible that the relevant formal object be the case but for the relevant emotion not to be elicited, such as when we are not aware of the formal property or when we judge neutrally that it obtains. These conditions on the characterisation of the formal object entail that it must not be something trivial like 'whatever property elicits emotion x for me'.[7]

In addition, and as already mentioned in Chapter 1, the concept of the formal object is the concept of a property that individuates different emotions and we can use it to identify individual emotions of a type. So, we also require that our characterisation of the formal object is such that the formal object can play these roles. As I discussed in Chapter 1, if the formal object is what individuates and what all instances of an emotion type share, then we can also use the formal object to assess an emotion for intelligibility, accuracy and appropriateness. These five roles often given to the formal object are thus:

(I) The formal object is what all instances of an emotion type share.

(II) The formal object is what individuates emotions.

(III) The formal object determines whether an emotion is intelligible.

(IV) The formal object determines whether an emotion is correct.

(V) The formal object determines whether an emotion is appropriate.

In this section, I begin in section 6.2.1 by fleshing out what we are committed to by accepting roles (I) and (II). This will lead to the objection that the formal object

[7] As Teroni puts the point, we should not understand the formal object in a way which makes it trivial, such as if being frightening, for instance, is 'the property of inducing fear in a given individual or species' (Teroni 2007, 397). This is because we need to allow that particular objects can 'exemplify formal objects independently of the occurrence of the relevant emotions (say for lack of relevant information)' and that the relevant emotions can 'occur without particular objects exemplifying the relevant formal objects' (397).

of an emotion is not what individuates. In responding to this objection in section 6.2.2, I argue that the concept of the formal object is best understood as the concept of a property which produces experiences of a certain kind in a responder, namely a response-dependent concept, and if we understand it in this way it does individuate. In section 6.2.3, I draw together my conclusions for 6.2 and how they affect our understanding of the formal object and the intentional nature of an emotion.

6.2.1 What the role of the formal object commits us to

Because of the role in (I), we are committed to the necessary conditions that:

A. If an emotion is of type E then it is the emotion it is in virtue of being elicited by something as f.

   - For example, if Joan is afraid, then her emotion is fear in virtue of being elicited by something as dangerous or as frightening.

B. If an emotion is not the emotion it is in virtue of being elicited by something as f, then it is not of type E.

   - For example, if Joan's emotion is not the emotion it is in virtue of being elicited by something as dangerous or as frightening, then it is not fear.

Explicating these conditions for (I) makes clear that we can be wrong about what emotion we are experiencing. For example, if I am driving in a dangerous or frightening part of town and experience an emotion, that emotion is fear because it is elicited in virtue of the area as dangerous or as frightening. I may, however, identify my emotion as anger with my innocuous passenger. If asked why I am angry with my passenger, I will try to show that the passenger is somehow being offensive; in other words, I will try to show that the passenger instantiates the formal object of anger and thereby make my emotion intelligible. It nevertheless remains the case that my emotion is not anger because it is not *elicited* by something as offensive; it is elicited by something as dangerous or as frightening.

What does not follow from (I) is that, if an emotion is the emotion it is in virtue of being elicited by something as f, then it is an emotion of type E. Nor does it follow that, if an emotion is not of type E, then it is not the emotion it is in virtue of being elicited by something as f. This, however, is how we might be tempted to explicate (II), the role that the formal object is what individuates emotions:

C. If an emotion is the emotion it is in virtue of being elicited by something as f, then it is an emotion of type E.

- For example, if Joan's emotion is the emotion it is in virtue of being elicited by something as dangerous or as frightening, then it is fear.

D. If an emotion is not of type E, then it is not the emotion it is in virtue of being elicited by something as f.

- For example, if Joan's emotion is not fear, then it is not elicited by something as dangerous or as frightening.

There are, however, apparent counterexamples to C and D. If they hold up, then it seems that emotions cannot be individuated on the basis of their formal object alone and we must deny (II). Doing so does not detract from (I). I return to (III), (IV) and (V) in section 6.3.

6.2.2 Counterexamples to the individuating role

William Lyons (1980) identifies the formal object of fear as the dangerous and gives the example of feeling excited in virtue of something as dangerous and being afraid in virtue of something as dangerous.[8] Let us grant for discussion that the formal object of fear is the property of being dangerous. It seems that both emotions are what they are in virtue of something as dangerous, but they are not the same emotion. Therefore, the objection goes, the formal object is not what individuates. Lyons' solution is to propose that not all cases of excitement are in virtue of something as dangerous whereas all cases of fear are; therefore, the dangerous is not the formal object of excitement (Lyons 1980, 101). This is quite right. The dangerous does not meet A or B, and therefore fails to meet (I), with regards to excitement.

Fabrice Teroni (2007), however, gives some different examples and concludes that the formal object is not what individuates emotions, thus rejecting (II). His examples are of thrill and fear, and of shame and contempt. He identifies the formal object of both thrill and fear as the frightening, and the formal object of both shame and contempt as the shameful.

Teroni proposes that thrill is only ever elicited by something as frightening, and that it is in virtue of the particular object as frightening that the emotion is thrill. If he is right, then A and B are met for thrill, but on the basis of the formal object alone, we cannot distinguish between thrill and fear. Similarly, both shame and contempt are only ever elicited by something as shameful.

I do not find these examples convincing for two reasons.

[8] I use 'in virtue of' as short-hand for 'in virtue of being elicited by' because Lyons does not talk about elicitation. This does not affect the point being made in this paragraph.

The first reason is that one emotion's simply being elicited by the same thing as another emotion does not entail that the formal object of the first emotion is the same as the formal object of the second. We can have beliefs about something as frightening or about something as shameful, beliefs which are elicited by something as frightening or something as shameful, respectively, without also experiencing an emotion. It is therefore possible to be in a state which is elicited by the same thing as an emotion without being in that emotion state. If it is possible for non-emotions to be elicited by something as frightening or shameful, then it may also be possible for another emotion to be elicited by something as frightening or as shameful, so long as the emotion has its own formal object *in virtue of which* it is the emotion it is. We thus need to look closer at *how* the emotion is elicited; and this gives rise to a question about the concept of the formal object and how the formal property relates to the emotion, rather than a question about the formal property itself. By returning to the difference between fear and excitement and starting with the assumption that the formal object of fear is partly the property of being dangerous, we can cash out how this works.

Suppose the free solo-climber Alex Honnold, who scales cliffs without any gear and by himself, feels excited by what he does. Even though Honnold is a highly trained and conditioned climber and the possibility of his falling is very low, what he does is dangerous. There is a fact of the matter that, as a human being, the cliff-face poses a threat to his well-being and is dangerous. Let us suppose further that, like many other adrenaline-fuelled extreme sports enthusiasts, he does not feel afraid; rather, he feels excited. And, crucially, he feels excited exactly because scaling great heights with no safety equipment is dangerous. That is, in virtue of the situation as dangerous, he is excited and his excitement is elicited by something as dangerous. We have already seen that the dangerous is not the formal object of excitement because it fails to meet (I); however, we need to ask: what is the formal object of excitement which does meet (I)?

We do not only feel excited in the face of danger: Honnold may be excited by going on holiday or seeing an old friend for coffee. This suggests that the formal object of excitement is more plausibly something like being presented with a tangible and immanent good. Whether the situation presents a tangible and immanent good depends on the features of the situation and what relation holds between the particular object of the emotion and the subject, just as whether there is a danger depends on features of the situation and what relation holds between the particular object and the subject. Honnold, for example, is a skilled climber and I am not. Even while the situation is dangerous for both of us, it also presents a tangible and

immanent good for Honnold in virtue of which his excitement is elicited, where the tangible and immanent good is something like the chance to meet a challenge. It does not present a tangible and immanent good for me.

This is a feature of the formal object as a relational property. A relational property is a property which holds between two relata. For example, the man who is the child of my parents is my brother. His being my brother is relational: he is my brother because of the relation that holds between him and me, that we are children of the same people. Similarly, being dangerous is a relational property because something is only dangerous if it stands in a particular relation to something for which it is dangerous. A cliff-face is dangerous for ground-bound humans because we cannot fly but it is not dangerous for eagles because they can fly. Because he is a human, the cliff-face is dangerous for Honnold. However, because he is a skilled climber, the cliff-face also presents him with a tangible and immanent good in virtue of which his excitement is elicited. We therefore see that the relational property is also relative to the individual. To distinguish between something which is dangerous and also the formal object of fear and something which is dangerous and also the formal object of excitement, let us talk about the formal object of fear as the frightening and the formal object of excitement as the exciting.[9]

Excitement, then, does not have the same formal object as fear. Teroni's version of the objection, however, was that thrill does. But a similar story can be given in response.

It is not clear to me that thrill is different to excitement, except that it is possibly the class of excitement which is elicited by something as frightening.[10] By definition, then, Teroni would be right that thrill is elicited by something as frightening. But if thrill is just the class of excitement which is elicited by something as frightening, then the formal object is rather the property of presenting a tangible and immanent good—and a tangible and immanent good is contingent on, in this case, the situation as frightening. Perhaps the tangible and immanent good arises out of showing off bravery, in which case being frightening directly elicits excitement. Or, perhaps the tangible and immanent good arises out of the way the situation elicits fear by being frightening, where the tangible and immanent good is the fearful

[9]This is a statement of terminology, not a statement about the formal object as response-dependent.

[10]Although even this is not clear to me. We talk of the 'thrill of the chase', for instance, but thrill here is not what it is in virtue of being elicited by something as frightening. Thrill is more likely the class of excitement elicited by something as dangerous but, as we have seen, simply being elicited by something as dangerous does not entail that it is elicited *in virtue of* of that thing as dangerous.

adrenaline.[11] In this latter case, excitement is elicited indirectly by something as frightening, via fear.

We therefore see that being elicited by something which may be the formal object of some other emotion does not entail that that the two emotions share a formal object. In other words, because, e.g. thrill, is elicited by something as frightening, directly or indirectly, it does not follow that it is *in virtue of* being elicited by something as frightening that it is thrill, except contingently.

The second reason why I do not find the examples convincing is that, if we look closer at why emotions of the same type share a formal object, that is, if we look closer at (I), we see that the concept of the formal object is broader than someone like Teroni allows.

If our concept of the formal object is of something which all instances of an emotion type share, we need to say more about how the formal property relates to the emotion in order to establish why a certain property elicits a certain emotion. After all, a proposition as true can be the object of believing, wishing, imagining, etc., but it is *in virtue of* being taken as true, rather than being imagined as true or wished it were true, that it is the object of belief. So, why are emotions of a certain type elicited by the same kinds of properties? It is in answering this question that we uncover how the concept of the formal object is a response-dependent concept, and we see that the formal object does individuate.

Imagine Scared-Honnold, Honnold's doppelgänger with the same abilities. Unlike Honnold, Scared-Honnold feels absolute terror while scaling cliffs and he only does it for the money. Both Honnold and Scared-Honnold are humans, and climbing cliff-faces is dangerous for humans. Both are also skilled climbers and climbing cliff-faces is a challenge they can meet and a dangerous situation—frightening, even—they can enjoy. For both of them, the situation presents a tangible and immanent good. In contrast, I am not a skilled climber. While the situation is dangerous for me as well, it does not also present a tangible and immanent good. Nevertheless, both scared-Honnold and I experience fear while Honnold experiences thrill. This is because our different cares and interests affect which property elicits an emotion in us.

Unlike Honnold, Scared-Honnold is not set on showing off his bravery or accumulating fame, although the situation is such that he could do both. Neither does

[11] Honnold's own description of his state of mind picks up on this fact that fear and excitement, or thrill, are so closely connected. 'I do get scared while soloing', he says. 'Something will happen and you have tht little jolt of *fear, excitement, or whatever you want to call it*' (Schmalz 2012, my emphasis).

Scared-Honnold have a thrill-seeking nature and take pleasure from the adrenaline rush of being afraid, although the situation is such that he could do so. In other words, Scared-Honnold does not experience the situation as frightening and also as presenting a tangible and immanent good like Honnold does, either directly or indirectly. And unlike Honnold, Scared-Honnold does not feel thrill but, like me, he feels afraid. So, because Scared-Honnold and I have the cares we have, including ones for our safety, the situation as frightening elicits fear *in virtue of* being frightening. But because Honnold has the cares he has, including ones for fame and adrenaline, the situation as frightening elicits excitement *in virtue of* being exciting, not in virtue of being frightening. If this is the case, then the subject's goals, what she values, and what she cares about play an important role in whether an emotion, and which one, is elicited. It is because of the similarity in which our cares are affected, or are seen as being affected, that all instance of fear, anger, etc., share a formal object.

Now, if you have certain cares, then you will be disposed to respond actively to the ways in which those cares are being affected, perhaps by wanting to and trying to protect the objects of those cares from harm, or viewing them favourably in deliberation. Of course, any one person may have many cares, some of which conflict, and some of which are more important than others. Sometimes a care may not be motivating or intense at a particular time for a range of reasons, such as being momentarily forgotten about, not being relevant, or even because the person is suffering debilitating apathy. It is therefore possible for someone to have certain cares, for them to be affected, for her to recognise or judge that they are being affected, and not, in the moment, actively care about this being the case. However, other things being equal, if someone has a set of cares, and one of them is being affected in a given situation and she is aware that it is, then she will *care*. That is, she will react actively, engagingly—in line with *something she cares about being affected*. Having a care is, other things being equal, caring. And the concept of caring is partly of being responsive.

We therefore see that in order for an emotion to be elicited *in virtue of* a some property, the subject has to care, and the concept of caring invokes the concept of responding. As a consequence, the concept of the formal object of an emotion is not just the concept of a relational property. It is the concept of a relational property that stands in relation to someone actively caring who experiences the property as salient to her cares. In other words, the concept of the formal object is a response-dependent concept, where a response-dependent concept is essentially a concept of a property which produces experiences of a certain type in a responder (such

as Johnston 1989 describes). Applying the concept involves thinking that such a response is required, in a way still to be established. For example, we might say that our concept of redness is essentially a concept of something that produces red-experiences. If something is red, then it would produce red experiences. Or, our concept of the frightening is essentially a concept of something that elicits fear. If something is frightening, then it would elicit fear.[12]

A response-dependent concept is distinct from a response-dependent property.[13] A response-dependent property is one which essentially stands in a certain relation to some kind of subjective response (Wedgwood 1997). We might think that red is an example of a response-dependent property. If something is red, then it essentially stands in a relation where it evokes a phenomenal red experience in normal perceivers in normal conditions.[14] We may, however, accept that the concept of red is response-dependent because it is essentially a concept of a property which produces red experiences, but deny that the property itself is response-dependent, that is, a property that essentially stands in a relation to a subjective response. Perhaps it is instead, essentially, a pattern of lightwaves.

Similarly, if the formal object of an emotion is a response-dependent property, then it would essentially elicit an emotional response of a certain kind in a class of perceivers in certain conditions.[15] For example, the property of being frightening, we might suppose, is one which essentially elicits fear in human perceivers in dangerous-for-humans conditions. But, while our concept of the frightening might essentially be the concept of something which produces fear experiences in us, the property of being frightening might also be independent of the kind of experiences that we have. It might simply be the relational property of being dangerous(-for-humans), or even the bearers of those relational properties, like the steep cliff-face.

We now also see that the emotions are what they are in virtue of being elicited by different properties, the frightening or the exciting, and so the counterexamples

[12]We could also understand a response-dependent concept as a concept which we acquire only through experiences of the relevant kind (Pettit 1991, 587). Without experiencing red, we will not have a full first-hand grasp of the concept 'red'; without experiencing fear, we will not have a first-hand grasp of the concept 'frightening'. I am not concerned with how we acquire the response-dependent concept and set aside this interpretation.

[13]Ralph Wedgwood (1997) and Jussi Haukioja (2013) both give useful overviews of response-dependent properties versus response-dependent concepts.

[14]Sometimes a response-dependent property is cashed out as a dispositional property, where it is a property that would invoke experiences of a certain kind in the right kind of subject in the right kind of conditions. Something can have a disposition without that disposition being made manifest.

[15]As Mikko Salmela (2011) points out, specifying the relevant class of perceivers and kind of conditions appears harder for emotion than for colour, and we need to be cautious of giving a trivial and highly subject-relative definition.

to the formal object as what individuates no longer hold. Teroni does have a further objection which targets this kind of proposal, however. His objection is that defining the formal object in the way I have bundles together a whole bunch of things, other than the formal object, which can be used to individuate emotions, such as action tendencies or other beliefs held by the subject. He writes that '[t]he worry here is that this might well lead to merely verbal distinctions, since it does not seem, for instance, that the "shameful" and the "contemptible" [or the "thrilling" and the "frightening"] refer to distinct values' (Teroni 2007, 402).

Granting that both shame and contempt are elicited by something as shameful and that both thrill and fear are elicited by something as frightening, I agree that they do not refer to distinct values. They do not refer to distinct values because they are responses to the same property. However, experiencing shame elicited by the shameful rather than contempt elicited by the shameful, or fear elicited by the frightening rather than thrill elicited by the frightening, does entail that there are distinct values on another level which change the way the shameful and the frightening are experienced. That level is determined by the subject's background, including her skills and cares. So, while the property that elicits different emotions may be distinct from other aspects which may be used to individuate the emotion, those other aspects still fall under the concept of the formal object because they elucidate *how* an emotion is elicited in virtue of the formal property. It is not a random bundling and the distinction is not merely verbal.

6.2.3 Clarification on the intentional nature of emotion

In this section, I have argued that the concept of the formal object is a response-dependent concept but I have not said much about what the property itself is. It is compatible with the view of the concept as response-dependent that one of its elements is an evaluative relational property like being dangerous, offensive, a loss, valuable to me, etc. Indeed, the concept of the formal object of an emotion does include that, whatever the particular object is, it stands in a certain kind of evaluative relation to the subject, like being dangerous, offensive, a loss, valuable to me, etc. I therefore accept the widespread view that the formal property is best understand as an evaluative relational property; nevertheless, our concept of it is response-dependent. This brings some consistency to the way people talk of the formal object, as highlighted at the beginning of the section. Those like Tappolet and Justin D'Arms are concerned primarily with the concept of the frightening, whereas someone like Prinz is concerned with the relational property of being dan-

gerous.[16] So, while the concept of the formal object groups together a number of elements, the evaluative relational property is one of these elements.

There is also now clarity concerning the role of the formal object in the intentional nature of an emotion. The formal object need not be mentioned in the explicit intentional content of an emotion. Both Honnold and Scared-Honnold experience different emotions which are about something as dangerous or about something as frightening. Their emotions, however, are elicited by different formal objects and the emotional experience as a whole represents the dangerous or the frightening *as* exciting, for Honnold's excitement; or the emotion as a whole represents the dangerous or the frightening *as* frightening, for Scared-Honnold's fear. It seems wrong, therefore, to talk of emotions as *perceptions* of value because the value does not enter explicitly into the intentional content. The value is rather represented by the emotional experience as a whole and, in this way, the subject can respond to something as dangerous or as frightening without necessarily recognising that it is so.[17] Nevertheless, the intentional content can still be expressed as 'x' *represented as frightening*, or, *the frightening x*. In this way, the formal object is both represented in the expression of the intentional content and describes the intentional nature.

6.3 Using the formal object

At the beginning of section 6.1, I introduced two conditions for the characterisation of the formal object if it is to enhance our understanding of an emotion. First, it must be possible to experience an emotion even when the relevant formal object is not the case or what elicits the emotion. Second, it must be possible that the relevant formal object be the case but for the relevant emotion not to be elicited,

If the formal object is a property which we characterise as response-dependent—

[16] Prinz in fact argues that the property of the formal object cannot be response-dependent. His argument is not successful because it relies on a misrepresentation of the contemporary response-dependent claim about properties. He defines a response-dependent property as 'one that would not exist without being represented as P by a human experience, judgement, or other mental state' (Prinz 2004, 61). From this, he seems to conclude that the property must be represented by a mental state at any one time in order to exist at that time. But, he argues, something can be a loss, a danger, etc., or another relational property, without being represented or experienced as such. If these are the formal objects, as Prinz claims they are, then the formal object is not a response-dependent property. Prinz's definition is crude, however. A response-dependent property is one which, if something has it, essentially stands in a certain relation to some kind of subjective response. But it does not follow that something has the property only when there is someone responding; that is, when it is the object of a mental state right now. It only follows that, if something has it, it would elicit the relevant response in someone who is in the right kind of relation.

[17] As I argued in response to John Deigh's (2010) objection to the perceptual model in Chapter 4. My response thus perhaps does not salvage the perceptual model.

although it need not itself be a response-dependent property, essentially—then we need to ensure that we can define what the formal object is in a non-trivial way. If we do not, then we could end up by defining the formal object as 'whatever property elicits emotion x in subject S'. For instance, a dispositional analysis of the property of being frightening as the property of having the disposition to elicit fear in certain responders in certain conditions will run into problems when we try to specify what those perceivers and conditions are. Amongst other things, there is a lack of uniformity in what is frightening to people (D'Arms, 2005, pp. 4-5) and we may end up defining the conditions as those where emotion x is elicited in subject S. But then we fail to meet the two conditions because it is not possible to be wrong about the formal object or for the formal object to be present but for the relevant emotion not to be elicited. We might object, that is, that on the view I am proposing, the view of the concept of the formal object as a response-dependent concept, these two conditions are not met and the formal object comes out trivial.

Luckily, plenty of insightful work, especially by Justin D'Arms and Daniel Jacobson, has been done on an alternative, what is often called 'neo-sentimentalism'.[18] According to the neo-sentimentalist, 'to apply a response-dependent concept ϕ (i.e. to think that X is ϕ)'—what we do when characterising the formal object of an emotion but not necessarily what we do when we respond to the formal property—'is to think it appropriate (merited, rational, justified, warranted) to feel an associated sentiment F toward X' (D'Arms 2005, 3). In this section, drawing on the work done by neo-sentimentalists like D'Arms and Jacobson, I argue that the two conditions are met by showing how the formal object can be used to assess whether an emotion is merited in terms of its intelligibility, correctness and appropriateness, the remaining three roles for the formal object.

6.3.1 The formal object in assessments of intelligibility

John is afraid of the snake. Can the formal object, characterised as response-dependent, fulfil the required role of determining whether his emotion intelligible? It can, and it can do so in two ways. Because it can be used, we can meet the first condition.

The first way is by determining whether the particular object seems a certain

[18]For representative work, see D'Arms (2005) and D'Arms and Jacobson (2000), but also John McDowell (1984) and David Wiggins (1987). Neo-sentimentalists are primarily concerned with the nature of evaluative properties, not so much with emotions (although D'Arms and D'Arms and Jacobson argue that examining the sentiment, the emotion, provides necessary elucidation of the property). Some are concerned with evaluative properties as a whole, while others focus on only a few, such as the shameful, the fearsome, the offensive.

way to the subject, regardless of whether it really is that way. If John's fear is not elicited in virtue of something as frightening *to him*, then it is not intelligible. Even if he denies finding the snake frightening, the very fact that he experiences fear casts doubt on his denial rather than casts doubt on his finding it frightening. If he genuinely did not find it frightening, consciously or not, then his emotion, as fear, is not intelligible. Further, even if the snake is not really frightening, his emotion is still elicited in virtue of it as frightening *to him* and is in that way made intelligible.

The second way is whether the emotion coheres with the subject's other currently held mental states. Suppose that John has a phobia of snakes and lives in London where he is not likely to have ever encountered, or ever to encounter, a dangerous snake. He knows his snakes and believes that the snake in front of him is not dangerous, nor does he believe that he ought to find it frightening. In other words, he believes that the situation does not merit fear. Despite this, he is afraid.

On the one hand, his emotion is intelligible in the first way because the snake is frightening to him, even if he finds it frightening regardless of his other beliefs. We do not think that he does not really find the snake frightening, that his fear is not genuine. Rather, we think that he *ought not* to feel fear because the snake is not dangerous and ought not to be found frightening—and this is exactly why we treat phobias as pathologies and try to fix them using techniques such as exposure or desensitisation.

On the other hand, however, his emotion is not entirely intelligible because, while he finds the snake frightening, he denies that it is dangerous. If John genuinely believes that the snake is not dangerous and genuinely does not believe that the snake *is* dangerous, then his finding it frightening is unintelligible because part of finding something frightening is also finding it dangerous. Perhaps John is in denial or unaware of the fact that he actually believes deep down that the snake is dangerous. Even so, his overall experience is not entirely intelligible because his fear is in conflict with what he actively takes himself to believe, or what he actively takes himself to believe is not what he really believes. His emotion might be what is problematic, but so might what he believes he believes, or what he genuinely believes. In order to determine which is problematic, we must establish which is getting things right, and that involves establishing whether the formal object of the emotion is as things are: dangerous and meriting fear. His overall experience, however, is not entirely intelligible and the concept of the formal object as response-dependent allows us to assess how his responses cohere with his other mental states and how he is subjectively rational.

We therefore see that the concept of the formal object as response-dependent

does play a role in determining whether an emotion is intelligible. The kind of response that is merited must be intelligible, and what is intelligible is independent of the subject's current emotional experience. Therefore, we meet the first condition which is that it is possible to experience an emotion even when the relevant formal property is not the case or what elicits that emotion and the concept of the formal object as response-dependent need not be trivial.

6.3.2 The formal object in assessments of accuracy

The formal object does also determine whether an emotion is accurate because it provides accuracy conditions for whether the emotion is fitting. Because it does so, the concept of the formal object as response-dependent meets both of the conditions.

We can say of an emotion that it is appropriate or inappropriate, which I shall address in section 6.3.3, but as D'Arms and Jacobson argue, 'there is a crucial distinction between the question of whether some emotion is the right way to feel, and whether that feeling gets it right' (D'Arms and Jacobson 2000, 66). Whether an emotional response is merited could thus be in terms of whether the particular object of the emotion fits the circumstances (D'Arms 2005, 4).

As I have argued, the concept of the formal object is both of a property and of how that property elicits an emotion through the subject as caring. We can therefore assess an emotion for accuracy and fittingness in two ways. We can ask whether the emotion fits because the particular object has the relevant property, and we can ask whether the emotion fits because it is elicited in the right kind of way. These are the two sets of questions about the very notion of the formal object I identified in section 6.1.

With regards to the property, for example, fear will be fitting if the particular object really is dangerous, independent of whether the subject finds it so. It is dangerous if it is a fact that the particular object really does stand in some kind of threatening relation to the subject. Fear will also be fitting because the situation is dangerous or frightening if the subject really has the relevant cares relative to which the situation would be dangerous, although this can be difficult to establish. Nevertheless, Julien Deonna (2006) makes useful headway in fleshing out what some conditions must be for whether the subject really has the cares, where the cares are facts about the subject.[19] He argues that there will be some species-relative facts, such as the fact that cliff-faces are dangerous to humans who cannot fly. He

[19]Salmela (2011) argues that Deonna's attempt at fleshing out the conditions is not successful and I will respond to his criticism in Chapter 7.

also argues that there will be individual-relative facts. These are composed of both long-standing values we might have, as well as our character traits or temperaments. For instance, Honnold has a long-standing goal to be one of the best climbers in the world. Because of this long-standing goal, scaling the cliff-face presents a tangible and immanent good and does merit excitement for him even though, as a human, it also merits fear. Scared-Honnold and I are both of cowardly characters. Scaling the cliff-face is thus dangerous for both of us, relative both to ourselves as human and relative to ourselves with the things we care about, and does merit fear. Given our characters and long-standing values, excitement is not a merited response and we would be wrong to feel excited.

This way of assessing accuracy and whether the property obtains is important because it allows room for a subject to react to the wrong cares, so to speak. If her emotion is in conflict with her character traits or long-standing evaluative tendencies, then it is not fitting. Suppose John is a world-famous snake scientist who has made his life's aim to share his love and knowledge of snakes. He is being filmed for television while interacting with a highly dangerous snake. He knows how to treat it, has done similar things in the past and has been looking forward to it. This time, however, he feels terrified. While it is the case that his fear is fitting in relation to himself as human, it is not fitting given who John is. This is not simply a matter of his fear not being intelligible; it also gets things wrong by going against John's own long-standing values to which he is committed.

We can also ask whether the emotion fits because it is elicited in the right kind of way. If I am one of the subject's in Schachter and Singer's experiments described in Chapter 4, for example, then my anger is not accurate because it is elicited by an induced chemical change together with a stooge acting angrily; it is not elicited by my caring and there being something offensive. Suppose that the stooge, in his angry antics, actually does something offensive without my knowing—perhaps he tears up my questionnaire behind my back. In such a case, my emotion is still not elicited in virtue of something offensive, even though it would be merited because there is something offensive present.[20]

All of this shows the ways in which we can establish what emotion is fitting in terms of the formal object, independent of the subject's emotional experience.

[20]It is less clear what to say about the case where I also observe the stooge's offensive act. Is my emotion elicited in the right kind of way? The formal object as a response-dependent concept, however, still helps to give conditions for establishing whether the emotion was elicited in the right kind of way, even if we cannot definitively say which was the case. If my anger was elicited by the adrenaline injection and interpretation of my feelings in the given context, then no. But if my anger was elicited by caring about my questionnaire being torn up, then yes.

We can thus, again, meet the first condition because it is possible to experience an emotion even when the relevant formal object is not what elicits the emotion, because the emotion is elicited in the wrong kind of way, or it is not the case because the emotion is not fitting. We can also meet the second condition because, given what evaluative relation holds and who the subject is, the situation does merit an emotion regardless of whether the subject experiences the emotion. That is, it is possible for the formal object to be the case but for the relevant emotion not be elicited.

6.3.3 The formal object in assessments of appropriateness

Finally, the concept of the formal object as response-dependent also determines whether an emotion is appropriate. One way in which an emotion could be appropriate is by being fitting, in the sense detailed above. An emotion can also be appropriate morally, prudentially, or in some other prescribed way. For example, we can ask: ought the subject to have the cares she has in the first place? A paedophile may feel angry when his favourite pay-per-view website is shut down. It is true that something he values has been removed to his detriment and that he apprehends the scenario as such, yet he really ought not to value child pornography. His emotion is accurate and intelligible but it is entirely inappropriate. Anger, in such a scenario, is not merited and not just anything will be offensive, thereby meeting the first condition that it must possible to experience an emotion even when the relevant formal object is not the case or what elicits the emotion. The subject may fail to respond to the formal object because she does not have the morally worthy cares or because she is unaware of the moral, prudential, or other reasons which determine what response is merited, thus meeting the second condition, that it must be possible that the relevant formal object be the case but for the relevant emotion not to be elicited.

6.3.4 A final example

Let us walk through an example as a final illustration of how the concept of the formal object as response-dependent is not trivial.

Suppose that Jesse, a good middle-class white woman, has an implicit attitude towards black people, taking them as dangerous.[21] Let us suppose that she formed this attitude simply by growing up in an all-white community which was not explicitly racist. However, she is now politically aware, believes in justice for all, and

[21]Thanks to Sophie Stammers and the King's College London Rationality and Reasoning reading group for raising and discussing this case with me.

would not endorse the underlying attitude if it were brought to her attention. In fact, she actively argues against racism and rejects outright generic statements like 'All x are y' which express problematic stereotypes. Despite all of this, she often finds herself feeling afraid when walking through majority-black Lewisham to attend her evening yoga class. In this example, her fear is elicited by the people around her as dangerous and she finds the situation frightening because of a problematic attitude she unknowingly has.

Jesse's fear is intelligible in that the situation is frightening to her. She does not endorse the attitude towards black people as dangerous and does not understand why she feels jumpy whenever in Lewisham at night, but she nevertheless finds the situation frightening. It is not intelligible in the wider sense of her subjective rationality, however, because it is in conflict with other beliefs that she has.

Her fear is not accurate. She experiences the situation as frightening because of her implicit attitude towards black people as dangerous. But it is not the case that all black people are dangerous or that the situation is dangerous. We might appeal to how growing up in a white society in a world where the media presents black people as criminals has led to her forming the attitude she has; doing so, however, does not make either her attitude or her emotion correct. It only makes it intelligible. It remains the case that her fear inaccurately represents what danger there is. Further, the situation is not frightening because it does not merit fear. Fear is not fitting because the situation is not dangerous, hence not frightening and therefore does not merit fear; but it is also morally inappropriate. It is potentially even prudentially inappropriate because her fear might cause offense and thereby create a truly dangerous and frightening situation.

Suppose Jesse's childhood and life experiences were completely different and that from a very young age she and her family had been subject to various crimes committed by people of a different ethnicity to her. Again, this does not make her implicit attitude correct, it only makes it intelligible and does not change the accuracy of her current emotion. Suppose now that it really is the case that Jesse is in a dangerous situation. If so, then her emotion accurately represents how things are even if how she comes to experience the emotion is based on false associations. This, however, is not a problem specific to emotions as similar complaints abound, for instance, with Gettier examples regarding true belief and knowledge.

Her fear is also not accurate based on what her own cares and interests are. She actively endorses anti-racism and in all other aspects of her life works against racism. Feeling fear goes against and inaccurately represents her own long-term values.

Now consider Jesse's racist twin, a white supremacist with explicit racist beliefs. The kinds of things she cares about and her background both make it the case that being stuck in a majority-black area is genuinely frightening *for her*. But how a situation seems to the subject pertains to an emotion's intelligiblity and we need not be forced into saying that fear is also merited and that the situation genuinely is frightening. On the one hand, the reasons why the situation is frightening include, amongst other things, the racist twin's explicit racist beliefs about black people, of which presumably many are false. So, while the situation is frightening *to her* it may not be dangerous and fear is thus not merited. On the other hand, we can also assess the evil twin's fear for appropriateness. Given that she feels fear on the basis of morally dubious commitments, fear is inappropriate.

6.4 THE EMOTIONAL REASONS

We can finally return to what reasons an emotion could provide access to in virtue of its intentional nature. I have argued that the concept of the formal object of an emotion is response-dependent and that it is implicitly represented in the emotion's intentional content. This can be expressed as 'the $FO\ x$', where x is the particular object. For example, John's fear is about 'the frightening snake'. I have also argued that one element to the concept of the formal object is a relational property on which the merited response is grounded. The property is relational because it holds between the agent and some feature of her environment but it is also relative to the agent as a member of a group or species, or as an individual. It is not relative, however, to how things seem to her. This reason can thus be expressed as 'the $RP\ x$', such as 'the dangerous(-to-me) snake'. Because the emotion is occurrent and elicited by some particular object as being a certain way, the emotion provides access to the potential fact that 'there is a $FO\ x$' or the potential fact that 'there is a $RP\ x$'.

Through looking at how an emotion can be assessed for intelligibility, accuracy and appropriateness, we saw that an emotion can be intelligible, accurate and appropriate in a variety of ways: in terms of how things seem to the subject, in terms of what response is merited, and in terms of how things are. In discussing intelligibility, I introduced the suggestion that an emotional experience is intelligible because of how things seem to the subject. We therefore get a third potential fact, that 'there is an $FO\ x$ (to me)'.

To illustrate, suppose that John backs away from the snake because he is afraid of the snake. Currently, 'he is afraid of the snake' functions as a placeholder for

whatever his reason is. We now see that there are three options for his reason:

1. How things seem to him: There is a frightening snake (to me).
2. How things are, in terms of what response is merited: There is a frightening snake.
3. How things are, in terms of what relation holds: There is a dangerous(-to-me) snake.

All of the options, 1-3, could be reasons. If the snake is frightening to John, then it is a fact about the experience that John is having and is therefore a reason. If the snake is frightening, then this is a fact about whether the snake really does merit fear. The fact may be true relative to different individuals or different value systems, but it is a fact nonetheless. Finally, if the snake is dangerous, then it too is a fact.

6.5 Conclusion

In this chapter, I examined the concept of the formal object and argued that it is best understood as a response-dependent concept. If we understand it in this way, then we can see how the formal object fulfils its five role. These were that it: (I) is what all instances of an emotion type share, (II) is what individuates emotions, (III) determines whether an emotion is intelligible, (IV) determines whether an emotion is correct, and (V) determines whether an emotion is appropriate. Such a characterisation of the formal object meets the two conditions that it is possible to experience an emotion even when the relevant formal object is not the case or is not what elicits the emotion, and that it is possible that the relevant formal object be the case but for the relevant emotion not to be elicited. The characterisation of the formal object as response-dependent is thus not trivial.

One of the implications of this understanding of the formal object is that it is only implicitly represented in an emotion's content; nevertheless, we can still identify three potential reasons which the emotional experience provides access to. For each of these potential reasons, we now need to ask two questions:

Q1. Can we act in light of the reason when acting solely on the basis of an emotion as an intentional state?

Q2. Can we act reasonably when acting solely on the basis of an emotional experience?

One might argue that the answer to both questions is ‘no’. With regards to the first question, we might object that even when we do act from the reason an emotion provides access to, we first have to believe that it is the case. We therefore do not act in light of the reason when acting on the basis of an emotion, but when acting on the basis of a belief. With the second, we might object that emotion just is the wrong kind of state to allow us to act reasonably from the emotion. In the next chapter, I respond to these two objections.

CHAPTER 7

ACTING REASONABLY IN LIGHT OF A REASON

'Empirical work on and common observation of the emotions tells us that our emotions sometimes key us to the presence of real and important reason-giving considerations,' writes Karen Jones (2003, 181), describing one of the key motivations behind the pro-emotion consensus, the school of views which rejects the opposition of emotion to reason. The reason-giving considerations, I argued in Chapter 6, arise from emotion's intentional nature and I articulated three potential reasons, as follows:

1. How things seem to the agent.
2. How things are, in terms of what response is merited.
3. How things are, in terms of what relation holds.

As discussed in the Introduction, this is not yet sufficient for an emotion to be involved in our conception of rational action. Even if emotions have a functional role that is not opposed to more traditional reason capacities, such as deliberation, and even if they do provide these reasons through their intentional nature, it does not follow that acting solely on the basis of an emotional experience is either acting reasonably or acting in light of a reason. That is, it does not follow that emotions have an important role in our conception of rational action, where part of our conception is that the agent guides her action in light of reasons and does not simply respond. I therefore need to address the two questions I posed at the end of Chapter 6 in order to show that acting on the basis of an emotion is acting in the right kind of way to be acting in light of a reason.

The two questions were these:

Q1. Can we act in light of the reason when acting solely on the basis of an emotion as an intentional state?[1]

Q2. Can we act reasonably when acting solely on the basis of an emotional experience?

In this chapter, I argue for positive answers to these two questions. I argue that we can both act in light of a reason when acting solely on the basis of an emotion and act reasonably when acting solely on the basis of an emotional experience. This contention requires responding to two brands of objections.

The first brand of objection targets what is required in order to act in light of a reason and is primarily concerned with acting rightly. The thought here is that, even if emotions provide access to reasons, we do not act in light of those reasons when acting solely on the basis of an emotion. In order to act in light of those reasons, we need to believe or judge that they are the case and it is through judging or believing that we act *in light of* them and guide our action. In section 7.1, I address this objection by responding to Sabine Döring's (2007; 2010) formulation of it.

The second brand of objection targets the kind of state that an emotion is and is primarily concerned with acting reasonably. There are two distinct objections I shall focus on. The first is that emotions are not evidence of the evaluative facts which are the reason or do not constitute justificatory reasons, as Michael Brady (2013) argues with regards to emotional experiences as putative reasons for belief. The second is that we are not entitled to believe or act on the basis of an emotional experience alone, as Mikko Salmela (2011) argues, also with regards to emotions as putative reasons for belief. In section 7.2, I address these two forms of objection and apply them to action.

In the process of exploring these objections, I will develop my positive proposal for how we can act both in light of a reason when acting solely on the basis of an emotion and act reasonably when acting solely on the basis of an emotional experience. I end, in section 7.3, by returning to the examples of Category E and examining how they fare, in terms of acting reasonably and acting in light of a reason.

[1] I take this first question to be about the principle of whether, when an emotion presents us with accurate information, we can act on the basis of that emotion. I therefore will use 'emotion' rather than 'emotional experience' when discussing the first question.

7.1 ACTING IN LIGHT OF A REASON WHEN ACTING ON THE BASIS OF AN EMOTION

As I argued in Chapter 5, to rationalise an action is to show that the agent acted in light of a reason. Acting in light of a reason requires having that reason in mind in some way, and guiding one's action by it. 'Having a reason in mind' is ambiguous. It might mean that the thing you have in mind is the reason itself, such as if it were a proposition. But it might also mean that we have in mind something that is represented and expressed by a proposition, but which is not itself a proposition, such as the emotional reasons I identified in Chapter 6. Nevertheless, if acting for a reason requires having it in mind, it is not enough if we simply track information that could constitute reasons; we need to be aware of them as considerations in favour of acting and respond in an appropriate kind of way. The driving idea of this first objection I will consider is that emotions are not the appropriate kind of way in which to have reasons in mind. So, acting solely on the basis of an emotion will not be acting for a reason in the right kind of way.

I shall examine Döring's (2007; 2010) version of the objection, which is the argument that, while emotions have a role to play in practical rationality, they do not themselves directly rationalise an action by providing the agent's reason for acting. Rather, she claims, they provide non-inferential justification for the belief (or judgement) that things are as the emotion represents them to be.[2] It is this corresponding belief which does the rationalising work.

I begin in section 7.1.1 by presenting Döring's argument and giving an initial assessment of it in terms of her background commitments. While her background commitments are problematic, we can nevertheless reformulate her argument so that her conclusion still follows. Thus, in section 7.1.2, I argue positively that we can act reasonably when acting on the basis of an emotion and, in section 7.1.3, I argue that we can also act in light of a reason when acting on the basis of an emotion.

7.1.1 DÖRING'S ARGUMENT

Döring's argument draws on two background claims. She endorses a very particular view of emotions as perceptions, what she calls 'affective perception', and she endorses a distinction between subjective and objective reasons.

Emotions, for Döring, are affective perceptions of evaluative facts. She develops this claim by laying out an analogy with perception, of which four commitments

[2] Döring uses belief and judgement interchangeably and I will too.

are important in her argument against the rationalising role of emotions. First, she is committed to the claim that emotions, like perceptions, have representational content subject to correctness conditions. If I feel fear when I spot a threatening shape hiding in the shadows late at night, my fear is a perception of and represents something dangerous to me. If the shape is simply a bush, then it is not dangerous and my fear does not accurately represent how things are. I discussed how emotions can be fitting and assessed for accuracy in this kind of way in Chapter 6, where I also identified three candidates for how the emotion could represent things to be. These were: x as dangerous(-to-me), x as frightening, or x as frightening (to me). Because Döring distinguishes between an emotion's representational content and its formal object, and thinks that 'an emotion's formal object need not enter into the emotion's representational content' (Döring 2003, 222), her focus is on the evaluative content made up of 'x as dangerous(-to-me)'. For this reason, I shall discuss the agent's reason as 'x as dangerous(-to-me)' or, more simply, 'x as dangerous'.

Second, she is committed to the claim that the representational content of an emotion appears correct to the agent: it definitely seems to me that there is something dangerous in the shadows, much like the Müller-Lyer lines definitely seem to me to be of different lengths. This gives us how things seem to the subject, such as 'frightening to me' or 'dangerous to me'.

I accept the first two commitments, but shall ultimately reject the next two. The third commitment is to the idea that, like perception, emotions do not involve regarding the content as true. It is not incoherent or contradictory, Döring argues, to feel fear despite knowing that there is nothing dangerous to be afraid of, much like I can continue to see the Müller-Lyer lines as different lengths even though I know that they are equal in length. Finally, Döring is committed to the claim that the contents of emotions do not stand in inferential relations with the contents of beliefs or with each other. To support this claim, Döring argues that we can have 'conflict without contradition' both with perception and emotion, suggesting that the contents do not stand in inferential relations (Döring 2008). If you believe that p and also believe that not-*(p & not-p)*, then you cannot also believe that *not-p*. With perception, in contrast, if I see the Müller-Lyer lines as unequal in length, it has no inferential bearing on whether I should not believe that they are equal in length. I can see them as unequal and believe that they are equal, without contradiction. With emotion, too, whether I experience a situation as dangerous has no inferential bearing on whether I also believe it is dangerous, as I can feel fear and believe that there is nothing dangerous without contradiction, Döring argues.

Another important background claim to her rejection of the direct rationalising

role of emotions, one that is independent of her commitments arising from her model of affective perception, is her claim that there is a distinction between subjective and objective reasons. I have already introduced the distinction—and rejected it—in Chapter 5. In quick summary, subjective reasons are 'what the agent justifiably sees as a reason because of the contents of his actual mental state' and objective reasons are what are reasons because of 'some objective features of the situation' (Döring 2010, 297).

Döring argues that subjective reasons are what rationalise action because they account for the 'phenomena of rational guidance', and are what allow us to meet the guidance condition for rational agency. The guidance condition is the condition that we comply with or are guided by our judgements about what is right (Döring 2007, 366). She refers to R. Jay Wallace:

> ...the motivations and actions of rational agents are guided by and responsive to their deliberative reflection about what they have reason to do. Unless this guidance condition (as we may call it) can be satisfied, we will not be able to make sense of the idea that persons are genuine agents, capable of determining what they shall do through the process of deliberation. (Wallace 1999, 219)

Döring takes the guidance condition to mean that 'we do what we do because we hold certain propositions true' (Döring 2007, 371) and argues that 'if emotions rationalise action, they do so only via the judgements they can non-inferentially justify' (387). This is because:

> ...an emotion can play a rational role only if its content is endorsed by the subject. In experiencing an emotion it appears to the subject that the world is as his emotion represents it to be, but he need not affirm the correctness of the representation. Therefore he must take his emotion's content at face-value, and that is: make a corresponding judgment, before his emotion can make it rational to act in a certain way. (Döring 2007, 383)

Before proceeding, we must be clear on the difference between rationalising and justifying.[3] Giving a rationalisation requires showing that the mental state or action was held or done in light of a reason. We rationalise a mental state or action by giving the agent's motivating reasons. A justification need not require that it was

[3]Döring herself does not explicitly make this distinction and at times seems to use the two interchangeably, but I think we should understand her as saying something like what I describe in this and the next paragraph.

held or done *in light of* a reason, only that it is true, correct or arose from a suitable mechanism, perhaps one that is reliable. Reasons which rationalise and reasons which justify may sometimes come apart. For example, someone with blindsight may reliably form true beliefs about what she does not consciously see. On a reliabilist theory of justification, her belief is justified and we can justify it with the reasons that her belief is true and that it arose from a reliable mechanism. If we want to rationalise the belief, however, we need to show that the agent held her belief in light of a reason, but we cannot do so because the agent does not have reasons in light of which she believes what she does.

An experience may not be the agent's motivating reason for belief, even while it does justify the belief. This is because an experience is not the kind of thing that can enter into logical relations. Indeed, experience provides non-inferential justification rather than inferential justification. So, if emotions and perceptions play a non-inferential role in the rationalisation of other mental states and action, it is because they provide non-inferential justification for the contents of the belief or judgement, which in turn is what would rationalise the mental state or action because it is the agent's reason.

With all of this in mind, we can construct Döring's argument like this:

P1 Taking the contents of one's state at face-value is to regard the content of one's state as true. [*given*]

P2 Emotions, like perceptions, do not involved regarding the content as true. [*affective perception*]

C1 Emotions, like perceptions, do not involve taking the contents of one's state at face-value. [*P1,P2*]

P3 Subjective reasons rationalise action. [*guidance condition*]

P4 Having a subjective reason requires taking the contents of one's state at face-value. [*subjective reasons*]

C2 Emotions are not subjective reasons. [*C1,P4*]

C3 Emotions do not rationalise action. [*P3,C2*]

Döring's argument relies on the crucial distinction between subjective and objective reasons. Her motivation for the necessity of subjective reasons to rationalise action is that they are required in order to meet the guidance condition for rational

agency, and she appeals to cases where it seems that our subjective reasons, what we see as reasons, come apart from what the objective reasons are. In these cases, she thinks that we guide our action and act rationally even when we do not act for an objective reason. Indeed, if we acted in line with our objective reasons, she argues, we could even be acting irrationally. If this is right, what rationalises our actions must be our subjective reasons.

As I have already discussed in Chapter 5, the example Döring gives to illustrate the necessity of the distinction is a case of being told by a doctor to take a certain medicine (Döring 2007, 287). The doctor is mistaken about the medicine and it will in fact harm you. In such a case, it seems that what you have most objective reason to do is not take the medicine because it will harm you. Nevertheless, it also seems that the doctor's advice provides you with a normative reason to take the medicine because you are justified in seeing it as a reason for action. If you do not take the medicine, Döring claims, you are acting irrationally even when what you have most objective reason to do is not take the medicine. This, presumably, is because you believe that you ought to take the medicine and yet do not take it, suggesting that what actually guides an action and makes it rational is that you are acting in line with your beliefs about what your reasons are. It is not that you are acting in line with what the objective reasons are. The reasons that rationalise an action are thus subjective reasons, Döring concludes.

While I do not think that the subjective-objective reason distinction is a good distinction to make, for reasons argued in Chapter 5, Döring's argument can be reformulated in a way that avoids the distinction. I shall briefly present two of the problems with the distinction as they apply in the current context of discussion, and then, in section 7.1.2, reformulate Döring's argument in a way which I believe uncovers the real force of her objection.

The first problem is that Döring's example does not show that the agent's subjective and objective reasons come apart or that the irrationality arises out of not acting according to her subjective reasons. Hence, the example does not motivate adopting the distinction.

In Chapter 5, I argued that the reason in light of which the agent acts is her motivating reason and that the agent can act in light of good and bad reasons. Not all objective reasons will be motivating reasons, simply because you need to be aware of a reason in order to act in light of it, and the agent's motivating reason need not be the same as what she has most reason to do. So, even if it is the case that the medicine will harm you and that is a *pro tanto* reason not to take it, if you are not aware that this is the case it cannot enter into your deliberations or be

your motivating reason. Döring erroneously conflates having an objective reason in light of which you act with acting for what you have most objective reason to do, something which talking instead of the agent's motivating reason helps to bring out.

With this in mind, it now becomes clear that the example does not show that subjective and objective reasons come apart, or that acting in light of an objective reason can be irrational. In the example as described, it is not clear that, if you do take the medicine, you are acting in light of any reason at all, even if there is a reason which does justify your action. Further, it seems that you are even acting against a *pro tanto* objective reason of which you are aware of as a reason, namely the fact that the doctor has told you to take the medicine. If you see taking the medicine as a good thing to do in light of what the doctor tells you, then you do have an objective reason. The doctor really did tell you to take the medicine; and, given the doctor's medical expertise, her telling you to take the medicine is a *pro tanto* reason in favour of taking it. Not taking the medicine is thus not acting for an objective reason of which you are aware and aware of as a reason and, as the example is described, not having a countervailing reason for why you so act. But now we can identify what the source of the irrationality is. It is not that you are acting against your subjective reasons or that acting for an objective reason can be irrational. Rather, you are acting against objective reasons of which you are aware of as reasons without having a countervailing reason for doing so.

Perhaps one could respond that the example in fact illustrates that the fact that the doctor told you to take the medicine is not an objective reason to take it. If so, then the irrationality cannot arise because you are acting against an objective reason—there is none—but because you are acting against a subjective reason. However, given that the fact seems to be a reason because of who the doctor is, a healthcare expert, it *is* a reason albeit one that is defeated in the situation by other reasons, in particular the fact that the doctor is, in this instance, wrong.

As a result, the example does not show that subjective and objective reasons come apart because the reason for which the agent ought to act is an objective reason. It also does not show that the seeming irrationality of the case arises from acting against a subjective reason, as the agent is acting against an objective reason of which she is aware, and is aware of as a reason. And it does not show that we can act irrationally when acting for an objective reason, because in order to act for that reason we need to be aware of it and be acting in light of it, in which case it is no different to Döring's subjective reasons. Nevertheless, I have not disagreed with the spirit of the distinction, so there needs to be a stronger argument against adopting it, an argument which can be found in the second problem.

The second problem is that the objective-subjective reason distinction cannot account for the normative force of reasons. If there is a *pro tanto* reason for you to ϕ, then you ought to ϕ. For example, suppose that the building is burning down. The fact that the building is burning down is a *pro tanto* reason for which you ought to get outside. But if you do not believe it is a reason, then you do not have a subjective reason. And if having a subjective reason is required for acting rationally, you do not have a reason to get outside unless you believe that you have one, on pain of irrationality. But you do have a *pro tanto* reason to get outside, which is the fact that the building is burning down. How do we adjudicate between what you have reason to do in this scenario? We cannot say that the *pro tanto* reason to get outside is defeated by your lack of belief that it is a reason, else we lose the normative force of reasons. At best it seems that we can say that there is a reason for you to believe that the building is on fire, namely the fact that the building is on fire, but that same fact is not a reason to act.

Because of the way in which the subjective-objective reason distinction cannot account for the normative force of reasons, and because Döring's example does not motivate adopting the distinction anyway, if there is a better option available which captures the same intuitive idea regarding acting in light of what we take to be the case, we should rather go with that option.

As I argued in Chapter 5, there is a better way to capture the intuitive idea behind the subjective-objective reason distinction, and that is to distinguish between subjective rationality and acting reasonably on the one hand, and objective rationality and acting rightly on the other hand. As I discussed in Chapter 5, one sense of acting rationally is acting in light of a reason where a reason is a fact that counts in favour of the action. So, if you act rationally, you are doing what you have reason to do. In this sense, acting rationally is acting rightly. Another sense of acting rationally is acting reasonably and need not be acting rightly or for a reason, but it is acting in a way that is warranted, considering the beliefs and desires that you have and the factors of which you are aware. Sometimes one can act reasonably but not rightly. In Chapter 5, I proposed that acting reasonably refers to our 'subjective rationality' and is acting coherently with our other attitudes: it is being subjectively rational. Acting rightly refers to our 'objective rationality' and is in line with the facts: it is being objectively rational. In good cases, we act both reasonably and rightly. We rationalise an action by identifying the agent's motivating reason, a fact, but we can still explain an action as reasonable by showing how she acted in line with what she took to be reasons. In this way, we can capture the intuitive idea behind the subjective-objective reason distinction. This alternative distinction between acting

reasonably and being subjectively rational, and acting rightly and being objectively rational, does not lose the normative force of reasons because it does not make the agent's motivating reasons of a different kind to normative reasons. Both are facts and, regardless of whether we believe that we have a reason, we have a reason to act.

Wallace originally introduced the guidance condition when discussing something like subjective rationality. He writes that what I am calling subjective rationality 'may be thought of as intentional activity that is guided by the agent's conception of what they have reason to do' (Wallace 1999, 217). On the basis of this, he proposes a 'guidance condition' for subjective rationality related to agency: 'the motivations and actions of rational agents are guided by and responsive to their deliberative reflection about what they have reason to do' (219).[4]

We therefore see that the subjective-objective reason distinction is not required in order to assess whether the guidance condition is being met because we have an alternative distinction, one that does not run into the same problem regarding normative force of reasons. Further, as argued in Chapter 5, what rationalises an action are the agent's motivating reasons. So, if subjective reasons are not required in order to rationalise action and are not required in order to act reasonably, then P3, 'Subjective reasons rationalise action', is false.

If P3 is false, then it might seem that P4, 'Having a subjective reason requires taking the content of one's state at face value', is irrelevant since it does not matter what having a subjective reason requires of us. P4, however, plays an important role in cashing out what the reason which rationalises an action requires of the agent, whether the agent needs to judge something to be the case or not. So, even if we reject that something like subjective reasons are required to rationalise an action, we might still want to say:

P4* Acting because of a reason requires taking the contents of one's state at face-value.

Adopting P4*, Döring's argument still goes through and this premise, I believe, is the real force behind her argument. There are two good reasons why we might suppose that P4* is *prima facie* true. First, in order to act because of a reason, we need to be aware of it and have it in mind as a reason. Having a motivating reason, after all, is having a reason in light of which you see an action as good. It is

[4]Wallace's particular focus is on how practical deliberation relates to motivation. I use his guidance principle to structure my subsequent discussion both because it is deliberately neutral between accounts of how we can be so guided and also because Döring centres her discussion on it.

through seeing an action as good that we guide what we do and meet the guidance condition. This may require seeing the reason as true, which is taking the contents of one's state at face-value. Second, a reason must be such that it can be used in reasoning. This may require that having a reason is being in a state whose contents stand in inferential relations, and being in such a state is taking things at face-value.

If emotions do not involve taking their contents at face-value and if the contents of emotions do not stand in logical relations, then it will follow that acting from an emotion is not acting for a reason. By addressing each of these considerations in the rest of this section, I shall flesh out how, exactly, we can act in light of a reason when acting on the basis of an emotion.

7.1.2 Meeting the guidance condition and acting reasonably

To refresh, the guidance condition, as Wallace defines it, is the condition that 'the motivations and actions of rational agents are guided by and responsive to their deliberative reflection about what they have reason to do'.

The guidance condition is most straightforwardly applicable to acting reasonably and being subjectively rational because it is a condition for acting coherently with our judgements about what to do. Suppose that you know that the medicine is harmful but also that your doctor is trustworthy: you are aware of reasons which motivate conflicting actions and need to make a judgement about what reason to act for. The guidance condition only requires that, if you judge that the fact that the doctor told you to take the medicine is the reason in light of which you ought to act, then you must respond appropriately. It does not say anything about which reasons you ought to act for or even what is a reason.

Even so, the guidance condition is still relevant to acting for a reason because, through deliberation and judgement, we aim to identify the reasons for which we ought to act in order to act rightly. Further, acting for a reason is having a motivating reason in light of which you act and is not simply responding to what reasons there may be: you must have a reason in mind in order to act in light of it. So, we can understand the guidance condition as applying to what we must do when we have a motivating reason in light of which we see some action as a good thing to do.

Not all acting in light of a reason will be the result of a deliberative judgement and the guidance condition does not require that it is. This is because it gives requirements for both deliberation and deliberative reflection. Suppose you do not deliberate about whether to take the medicine or not, simply because it does not occur to you to doubt the doctor's advice. You just take the medicine as she hands it

over. Afterwards, lying prone in hospital having your stomach pumped, you could still respond to questions about your action by showing why you so acted, given your hypothetical deliberations. You might add that, had you known the medicine was harmful, you would have acted otherwise. Being able to explain your action in this way is showing how you acted in line with what you had reason to do, given your hypothetical deliberations and what you were aware of at the time. Indeed, Döring acknowledges this hypothetical nature of meeting the guidance condition. She writes:

> ...in order to be rational, an agent must satisfy the condition that *he would so authorize the action*, were he asked to do so ...a motive for an action such as an emotion must at least be *hypothetically authorized by the agent's deliberative reflection* about what he has reason to do. (Döring 2010, 295, my emphases)

We thus see that the guidance condition can be met in two ways, either by deliberating and acting in line with one's judgement, or by being able to justify one's action by appealing to what one would have hypothetically judged to be a reason for acting. The latter requires that the agent be aware of the reason as counting in favour of acting, and that she is able to offer it as a reason that justifies the action because she would have hypothetically judged it to be a reason in favour of acting.

Clearly, acting purely on the basis of an emotion is not acting as the result of deliberation and such action will not meet the guidance condition in the first way. However, acting on the basis of an emotion, I shall now argue, can meet the two further conditions for hypothetical deliberation, without a mediating belief.

With regards to the first condition, can one have in mind the reason as counting in favour of acting at the time of acting if one acts on the basis of an emotion? It is not enough that one simply entertain a thought that happens to be true, like if we were imagining or wishing that things were so. In order for it to be the agent's reason that counts in favour of acting, she needs to be committed to its truth in some way. Döring takes this to mean that we must take the contents of our state at face-value, or judge that it is true. Emotions, she claims, do not involve taking their contents at face-value because it is not incoherent or contradictory to feel fear despite believing that there is nothing dangerous or frightening to be afraid of. However, in Chapter 6 I argued that one way in which an emotional experience could fail to be fully intelligible is if it does not cohere with the agent's other currently held beliefs and mental states. After all, we do think that you ought not to feel fear if you do not believe that something is frightening, indicating that, while there is not

outright contradiction or incoherence, your experience is still in tension with your belief.

Further, as I also argued in Chapter 6, the concept of the formal object of an occurrent emotion is a response-dependent concept with one of its elements an evaluative relational property and another of its elements the subject as actively caring. Together, the formal property, such as the frightening, elicits the relevant emotion. I argued that the concept of caring is of something that is active and responsive to the things one cares about being affected. If someone does not respond to her cares being affected, it is because she no longer cares, is not aware of how she is being affected, or is apathetic and not moved by the way the things she is supposed to care about are being affected. So, if the subject experiences an emotion, it is not enough for her to see her interests as being affected by the things in her environment; she has to care that she is being affected in that way.

The concept of caring thus involves more than having a list of abstract cares which may or may not be affected at any given time. If one actively cares, then one experiences those cares as being affected but, if one experiences those cares as being affected, then in virtue of actively caring, one is committed to the way in which they seem to be affected. I can be in other states which are not truth-attitudes but which have the same kind of content as what an emotion represents, such as imagining or supposing that there is a dangerous or frightening snake, but the nature of these states is such that I am not committed to the content in the way that I am committed to there being a dangerous or frightening snake if I experience fear. What this means is that an emotional experience is partly a commitment to the way things appear through having a caring attitude. While such an attitude is not a truth-attitude and so we do not run into incoherence or contradiction, we can still pinpoint the tension as arising between actively caring and denying that one's cares are being affected.

Finally, if one actively cares about something, then one will see some action addressing how the care is being affected as good in light of that care as being affected. So, while this emotional commitment is not a truth-attitude like belief, if the emotion is accurately representing some evaluative information, we are aware of and have in mind a reason in light of which we could see some action as good and in light of which we could act.

With regards to the second condition, can we offer the emotional reason as a justification for our action because we would have hypothetically judged it to be a reason for acting?

What form an emotional response can take varies. A fear response may be a

physiological reaction. It may be an action tendency. And, similar to an action tendency, it may be a disposition to believe that one is being affected in the way the emotion represents. You may not currently believe or believe explicitly that things are as they appear, because emotions are not constituted by beliefs or judgements. Nevertheless, in the case of an occurrent conscious emotion, where you respond quickly and intuitively without thereby judging that such-and-such is the case, part of that response includes a disposition to believe that things are as they appear because of your emotional commitment through caring.

The claim I am making is not that you have a *dispositional belief* that things are as they appear, where a dispositional belief is something like Harry's dispositional belief that plaid ties are hideous (Schwitzgebel 2014). Harry's belief is not often occurrent and at the front of his mind, but it remains currently true of him that he believes that plaid ties are hideous. In contrast, even if John is afraid of the snake, it may be currently false of him that he believes that the snake is dangerous. Nevertheless, he is disposed to form the belief, other things being equal, in a way that he is not disposed if he were to imagine or wish that the snake is dangerous. Unlike imagining or wishing, he is committed to the putative truth that the snake is dangerous because he experiences it as dangerous through caring about his well-being.

We should also not understand my claim as the claim that emotions involve an *implicit belief* that things are as they appear. On one interpretation of implicit beliefs, an implicit belief is a belief that can be swiftly derived from what one explicitly believes (see, for example, Dennett 1987). On another interpretation, most often discussed in the context of implicit biases and problematic attitudes such as racist or sexist attitudes, an implicit belief is a belief that one does not explicitly endorse but which influences one's behaviour and beliefs.[5] Such implicit biases, however, are not clearly beliefs (Gendler 2008; Schwitzgebel 2010). Nevertheless, an implicit belief on this latter interpretation is still some kind of implicit attitude we have towards making certain associations and, because of those associations, a disposition to behave in ways in line with those associations. We can understand my claim similarly as the claim that an occurrent conscious emotion is an attitude that takes things to have certain import and, because of that import, a disposition to believe that things are as they appear.[6]

[5]For an overview of evidence in support of and a defence of the existence of such implicit biases, see Jost et al. (2009).

[6]I am not committed to saying that it is permissible to form beliefs on the basis of an emotion, alone. I will discuss this matter in more depth in section 7.2.1.

If an occurrent conscious emotion involves a disposition to believe that things are as they appear, then we are in a position where the emotional reason is available to us to offer as a reason to justify our action. This is because we could have hypothetically judged it to be a reason for acting, seeing as we were emotionally committed to the way things were. So long as we are aware of the emotional content and would hypothetically endorse it, acting from an emotion can meet the guidance condition.

7.1.3 Acting in light of a reason

The guidance condition applies primarily to acting reasonably, not acting in light of a reason. Even if the guidance condition is met, I have not yet shown that acting from an emotion is acting for a reason rather than acting reasonably. More specifically, even if I am right about the way in which we can guide our action via emotions, I still have not shown that the contents of emotions can stand in inferential relations with the contents of beliefs, which is a requirement for something to be a reason.

The exact problem is as follows. Reasons are facts. As discussed in Chapter 5, a requirement we might have of reasons is that they can be used in deliberation and reasoning, and that they stand in inferential relations with each other. Because of this, we might endorse a view of reasons as true propositions expressing facts. If that is the case and emotions do not necessarily have propositional content, then emotions will not provide reasons even though they provide information in a non-propositional form. The arguments of the previous subsection, however, can be used to respond to this worry, together with a clarification of what reasons are.

Starting with the clarification, it is important to note that interpreting reasons as true propositions that express facts is not the only way to interpret reasons as facts. If reasons are rather what the true propositions represent—the truth conditions for the propositions—then another form of representation could represent the same reasons.

Here is why we might favour this understanding of reasons. In Chapter 5, I argued that not just any proposition can be a reason and argued against interpreting reasons as propositions which, if true, would favour action. The propositions need to be true. In Chapter 5 and also briefly in section 7.1.1 of this chapter, I argued against the subjective-objective reason distinction on the basis that it cannot account for the normative force of reasons. In order to account for the normative force of reasons and in order to allow that normative and motivating reasons are the same kind of thing, a proposition that is a reason must be true. These two arguments show

that, simply in virtue of being a proposition, something is not a reason. Rather, it is crucial that the proposition is true; that is, that the truth conditions for the proposition are met. While simply being true is also not sufficient for a reason to be an agent's reason—an agent will not act for a reason unless she has that reason in mind and acts in light of it—the insufficiency of being true is not due to the nature of reasons but due to the requirements of being able to act for it. *Prima facie*, it therefore seems that it is necessary for things to be the case in order for something to be a reason but it may not be necessary for the way things are to be expressed as a proposition in order to be a reason.

It may be necessary for a fact to be expressed as a proposition in order to be a reason because a reason still needs to be something (i) that we can be aware of and (ii) that can be a premise in reasoning. So, with regards to (i), can we be aware of a fact without adopting a propositional attitude? Consider an analogy. Tree rings represent the age of a tree but they do not do so by forming a proposition. Nevertheless, the same thing that the rings represent, the age of the tree, will be the truth conditions for the proposition stating the age of the tree. In other words, the tree rings represent the same fact as the proposition. Similarly, emotions represent how things are but not by a proposition. The same thing that emotions represent, such as something as dangerous, something as frightening, or something as frightening to me, will be the truth conditions for the proposition stating the fact that something is dangerous, that is is frightening, or that it is frightening to me. In other words, an emotion can represent a fact, which in turn can be stated in a proposition, and through an emotion we can be aware of a fact without adopting a propositional attitude.

Still, (ii) remains: reasons are the kind of thing that can be premises in reasoning. We might understandably be hesitant to accept that an emotional representation can be used as premises in reasoning. Just like tree rings cannot be premises, we might suppose that neither can an emotional experience or an emotional representation. As I argued in Chapter 5, mental states, and this would include emotions, cannot be reasons and so the emotion itself, as a state which is elicited by and represents the formal object, cannot be a reason. And, while I have argued that emotions involve a commitment to how things are, I have not shown that the contents of emotions stand in inferential relations with either beliefs or with each other.

However, this is where my arguments in the previous subsection can be called on. Importantly, whether the content of some state will stand in inferential relations with the content of another state is a function of what the state is. And, we see, emotions are states where we are committed to the content. In the previous sub-

section, I argued that feeling fear and believing that something is not frightening are in tension. While perhaps not a truth-attitude, we are committed to the way things appear through caring. Further, the way in which emotions are assessable for correctness does suggest that emotions involve a commitment to how things are, else it would not matter if they are accurate or not. Lastly, as I argued in Chapter 6 and have already called on in section 7.1.2, we do think that someone who does not believe that a situation is dangerous ought not to feel fear, and we can assess an emotional experience for intelligibility by how it coheres with the agent's other currently held mental states. If there were no inferential relation between an emotional representation and the content of a belief in virtue of the kinds of states they are, then it would not make sense to do either of these things. Taken together, an emotional representation can stand in inferential relations with the contents of other mental states like belief.

An emotional representation can also stand in inferential relations with other emotional representations, in virtue of the kind of state it is. Bennett Helm (2001, 2009) argues that the contents of emotions do stand in inferential relations with each other, and I find the basic idea compelling. For example, if I love my parents, it is because they have import for me; that is, I care about them, and it is because I care about them that I feel love. Because loving something requires that the object has import for the subject, if someone genuinely loves something, then that person will experience a range of emotions depending on how the thing with import is being affected. If I love my parents and they are harmed, then something that has import for me and which I care about is harmed and I will feel sadness; if I love my parents and they achieve some goal, then something that has import for me and which I care about has done something impressive and I will feel elation. If I did not experience emotions that are in accord with something having import for me, then we can legitimately question if I really do love what I say I love.[7] Indeed, it is often through our emotions that we realise that certain things have import for us. For example, on not feeling sad when a formerly close friend cuts ties, you realise that you no longer love her. The relations between different emotions, I contend, arise out of what the concept of caring involves, namely being active and responsive. So, if things have import and you care, then you will be responsive in ways appropriate to protect those cares, entailing that the content, to which you are committed, will stand in inferential relations.

[7]Helm argues that it is a rational failure not to experience the associated emotions. This stronger claim is not necessary for the basic point I wish to make, which is that the emotional representations do stand in basic inferential relations with each other.

Further, I argued in Chapter 1 that an emotional attitude, while not necessarily a disposition to experience the correlative occurrent emotion, is a disposition to experience other emotions and to have certain patterns of thought. In Chapter 3, I argued that the explanations in Category B turned on our conception of emotions as intentional and as having associated beliefs and desires. If emotional representations did not stand in some kind of inferential relation to each other and to the content of the beliefs we hold, then this would not be possible.

Emotional representations can stand in inferential relations, but can we use them in reasoning? My second argument in the previous subsection was that emotions involve a disposition to believe that things are as the emotion represents them to be. If so, what emotions represent can be expressed as a proposition and used in reasoning. If one is not aware of what an emotion represents or aware of it as a consideration in light of which one can act, one will not form a belief about it and offer it as a reason to justify one's action. But that is not the same as not being able to do so, if one is aware.

We therefore see that the concept of the formal object as response-dependent not only allows us to individuate between emotions, and assess them for intelligibility, accuracy and appropriateness, as discussed in Chapter 6, it also allows us to examine the role of emotions as providing potential reasons for action. Because we actively care and respond, we are committed to the way things seem to us in a way which enables us to act in light of considerations to which an emotion gives us access. And because we are committed, emotionally, to the way things seem to us, acting for those considerations can be acting in light of a reason when we act on the basis of an emotion. We therefore see that we can meet the guidance condition when acting solely on the basis of an emotion, and that we can act in light of a reason when acting solely on the basis of an emotion without a belief.

7.2 ACTING REASONABLY ON THE BASIS OF AN EMOTIONAL EXPERIENCE

Even if we could, in principle, act in light of a reason and guide our action when acting on the basis of an emotion, it does not follow that we are entitled to act solely on the basis of an emotional experience. This is because we may fail to be good rational agents. Perhaps the emotional experience is not suitably linked to the reason for which we act, some evaluative fact. For instance, we are entitled to act on the basis of perceptual experiences because perceptual experiences are evidence

for the perceptual fact and constitute sufficient justificatory reason for assenting to the perceptual fact.[8] Emotional experiences, on the other hand, may lack this kind of connection. They may fail to be evidence for the evaluative fact and they may fail to constitute a sufficient justificatory reason for committing to that fact through action. This is the first problem I shall consider in section 7.2.1.

But even if emotional experiences are evidence or constitute reasons, perhaps we still are not entitled to act on the basis of them because we are unable to give the conditions for when an emotional experience would be a defeasible reason and when not. This is the second problem I shall consider in section 7.2.2.

Both problems have been raised as challenges to perceptual theorists who argue that emotion has a similar epistemic role to perception in constituting reasons for, or justifying, belief. Either an emotional experience is not evidence for an evaluative fact, and an emotion cannot constitute a reason for believing that fact, which is what Brady (2013) argues. Or, we are do not have *prima facie* entitlement to assent to a belief about the evaluative fact because we cannot specify the defeasibility conditions, which is what Salmela (2011) argues.

In their arguments, both Brady and Salmela engage with perceptual theorists who argue that an emotional experience *by itself* constitutes sufficient reason *for a belief*, or that an emotional experience *by itself* is a perception of an evaluative fact about which we are justified in forming *beliefs*.[9] I am not necessarily endorsing a perceptual theory of emotion according to which emotions by themselves are perceptions of evaluative facts or constitute *sufficient* reasons—after all, my main motivation for adopting a perceptual approach was its minimal cognitive commitments and to use it as a working theory—and my focus is on action and not belief. Nevertheless, these two objections are pertinent to my project. Like the perceptual theorists, I am defending a role for emotion where acting on the basis of an emotion could be acting in light of a reason, which requires that acting on the basis of an emotion is not itself fundamentally problematic. If emotional experiences cannot constitute reasons for belief, sufficient or not, or if we are not entitled to assent to

[8] I shall clarify the use of 'justificatory reason' as well as 'sufficient reason' in section 7.2.1.

[9] Döring argues that 'in the rationalisation and justification of other states and actions the emotions play the role of perception, and in that sense, so I claim, they are perceptions' (Döring 2007, 376). Christine Tappolet writes, '…emotions are perceptions of values. Thus, fear would consist in perceiving something as fearsome…' (Tappolet 2012, 206-7). Julien Deonna aims to defend the idea that 'emotion, like perception, is a sensory means of gathering information about our environment' (Deonna 2006, 29) and Catherine Elgin argues that '[l]ike perceptions, [emotions] afford epistemic access to a range of response-dependent properties, such as being admirable or contemptible, and provide evidence of response-independent properties that trigger them' (Elgin 2008, 33). These theorists argue that an emotion, itself, is the perception of the evaluative fact and is sufficient to justify a belief, given the right circumstances.

a belief on the basis of an emotional experience, then why would emotional experiences have a different status in acting? Even though Brady's and Salmela's arguments are not designed to apply to action, they do provide a foil to examine the validity of acting on the basis of an emotional experience.

7.2.1 Emotional experiences as evidence and reasons

Brady (2013) argues that 'emotional experiences fail to constitute reason or evidence of any kind for evaluative judgements' (Brady 2013, 68). According to him, rather than constituting reasons that 'silence the call for justification' of a belief, they demand further reasons and are responses to reasons, not reasons themselves. In this subsection, I begin by clarifying terminology before seeing how the ideas behind Brady's argument could apply to action. In responding to his argument, I emphasise the difference between reasons and evidence, and argue that an emotional experience can be evidence for an evaluative fact and can constitute a justificatory reason for action.

Terminology

In section 7.1.1, I distinguished between justifying and rationalising an action or a belief. Rationalising requires showing that the agent acted (or believed)[10] in light of a reason. Justifying, on the other hand, need not require showing that the agent acted (or believed) in light of a reason. A belief, for instance, may be justified so long as it is true, or arises out of a reliable belief-forming mechanism, or it is suitably connected to what it is about, whatever that 'suitable' connection pans out to be. In this way, experience can non-inferentially justify a belief because of the kind of thing it is, without the agent inferring anything from it. In epistemology, it is this kind of justifying role that perceptual experiences are often supposed to have[11] and which perceptual theorists, like Döring, argue that emotional experiences share. When Brady talks of an emotional experience as a reason, we should therefore understand the relevant reason-role under discussion as what I shall call here a 'justificatory reason', rather than as the agent's motivating reason. So, by arguing that emotions do not constitute reasons for evaluative judgements, Brady is not addressing whether they can be the agent's reasons but whether they can justify a judgement or belief about the evaluative fact. When we apply these different

[10] I have not specifically addressed belief but I take the use of 'rationalising' and 'justifying' to be the same for both belief and action.

[11] For example, in debates about the conceptual content of experience, is is often assumed that experiences do provide reasons for belief but are not themselves had for reasons.

reason roles to action, we can distinguish between the agent's motivating reason and other reasons which might justify, such as her acting on the basis of a suitable sort of experience.

A second bit of clarification concerns reasons and evidence. Evidence is evidence *of* something's being the case or of some event; or it is evidence *for* the claim or fact that something is a certain way or that some event occurs. For instance, hoof-prints on a forest path are evidence of a deer's having been here or evidence of the passing of a deer; they are evidence for the claim or fact that there was a deer here or that a deer passed through. A reason is a reason *for* or *in favour of* acting, believing (or having the belief), supposing, etc., or a reason *why* things are a certain way. What something is evidence of or for need not be the same as what it is a reason for, in favour of or why. For instance, the fact that there is a bloody knife lying on the counter may be evidence of a murder having taken place or for the fact that there has been a murder. It is not a reason, however, for or in favour of a murder taking place, or a reason why a murder has taken place. It is rather a reason to think, believe, imagine, assert the claim, etc., that a murder has taken place, or a reason to act, such as to begin an investigation into the source of the bloody knife. If it turns out that the bloody knife is the butcher's knife at the end of a busy day, then the fact that there is a bloody knife lying on the counter is not, strictly speaking, evidence for the fact that a murder has taken place although it may nevertheless be a consideration in favour of making such a claim. But the fact that there is a bloody knife lying on the counter might still be a reason, albeit a defeasible one, to believe, etc., that a murder has taken place or to investigate the knife's source. Evidence and reason are connected, however. If something is evidence, then it will often also be a reason *because* it is evidence.

The claim that emotional experiences 'do not constitute reasons or evidence for evaluative judgements' thus could mean two things in the context of the current discussion. It could mean that an emotional experience is not evidence of there being an evaluative fact or for the evaluative fact. It could also mean that an emotional experience is not a justificatory reason for an evaluative judgement.

THE ARGUMENTS

The first argument which Brady puts forward is that emotional experiences, unlike perceptual experiences, are not sufficient reason for belief. A sufficient reason is one that 'is good enough to justify some action or belief by itself; a sufficient reason makes action or belief *permissible* for the subject, we might say' (Brady 2013, 84).

Sufficient reasons do not make the belief or action obligatory, nor do they entail that the action is correct or that the belief is true.

Brady compares a perceptual experience to an emotional experience. In normal circumstances and absent any defeaters, my perceptual experience of a red car parked outside constitutes good evidence of there being a red car parked outside or for the fact that there is a red car parked outside. We are permitted to form beliefs about the red car on the basis of such an experience, suggesting 'that perceptual experiences are sufficient reasons for beliefs about external objects and events' (Brady 2013, 86). The same cannot be said for emotion. Feeling fear on hearing strange noises in the house, for example, typically motivates us to seek additional evidence of and reasons for believing that there is danger. The emotional experience, alone, is not a sufficient reason for the belief. 'We do not, in other words, think that the demand for justification is silenced, or take the representational content of our emotional experience as true by default', Brady writes (87). We require more justification. Further, in most normal circumstances and with most emotions, we also think that it is impermissible to form a belief solely on the basis of an emotion. It therefore seems that emotional experiences at the very least have a weaker justificatory strength than perceptual experiences (89), and it is wrong to suppose that the same justificatory story can be given for both.

In discussing the differences between emotion and perception, Brady talks about emotional experiences as both evidence and reason. We should, however, treat them separately because something can be evidence of but not a reason for one thing, or a reason for but not evidence of another. If we treat them separately, three important points emerge that undermine the strength of the conclusions that can be drawn from the argument. First, even if an emotional experience does not silence the call for justification for belief, it does not follow that the emotional experience is not evidence for an evaluative fact. Second, evidence that constitutes a reason that does not silence the call for justification for one thing may silence the call for justification for another. Third, it then follows that, if an emotional experience is evidence for an evaluative fact then, independent of whether it silences the call for justification for a belief, it may still silence the call for justification for action; that is, it may constitute a justificatory reason for action. I shall develop each of these points in turn.

First, I agree that an emotional experience does not always silence the call for justification in the case of belief in the same way that a perceptual experience does. There are, however, two reasons why we cannot conclude from this alone that the emotional experience is not evidence for the evaluative fact. The first is that the

examples Brady uses, such as listening for further evidence that there is a prowler in the house, involve negative emotions like fear. But emotional experiences of more positive emotions often *do* silence the call for justification. An emotional experience of love, for instance, can silence the call for justification for all manner of beliefs and actions.[12] This suggests an alternative explanation for why, when hearing noises in the house and experiencing fear, we search for further reasons. The search for further reasons may not so much be answering a call for justification, as responding to a psychological need to prove oneself wrong or to get oneself out of the situation in any way possible—and one way is to establish that one is wrong. Feeling love or another positive emotion, in contrast, is not something we want to prove ourselves wrong about or a situation we want to get out of, but it is one we want to encourage, if we do find ourselves searching for reasons.

The second reason why we cannot conclude that the emotional experience is not evidence for the evaluative belief is that, while evidence can constitute reasons that silence the call for justification, evidence comes in different degrees and will not always constitute reasons that *silence* the call for justification. Take the use of evidence in the sciences, for example. In epidemiology, statistical correlation is often evidence for the fact that there is a causal link between some potential risk factor and some disease. But, so the mantra goes, correlation is not causation. There are many confounders, often unknown, meaning that the correlation alone is not enough to show that there is a real causal link. It nonetheless remains evidence for the fact that there is a causal connection, just weak evidence, and it does not constitute a sufficient reason by itself to silence the call for justification for the belief that there is a causal connection.[13]

Second, something that does not silence the call for justification for one thing may still silence the call for justification for another thing. The statistical correlation alone, for instance, is not strong evidence for the fact that there is a causal connection between a potential risk factor and some disease, and it does not constitute sufficient justificatory reason for believing that there is such a connection. It does, however, constitute a sufficient reason for believing that further study should be done. What this suggests is that, if a reason does silence the call for justification,

[12]For instance, Bernard Williams (1981) argues that a moral agent need not be thinking about what is morally justified and remain impartial all of the time; she can act on an emotion like love. Thinking about what is morally justified can sometimes be 'one thought too many'. Susan Wolf (2012) defends an interpretation of Williams' 'one thought too many' in which situations of love may even lie beyond justification.

[13]Thank you to Katherine Furman of the London School of Economics for discussing the use of evidence in health policy with me.

its doing so will be dependent on what it is being used for and in what context or, in other words, what it constitutes a reason for. Therefore, like for the first reason above, we see that if emotional experiences do not constitute sufficient reason for a belief, it does not follow either that they are not evidence for an evaluative fact or evidence of things being a certain way, or that they do not constitute sufficient reason for something else.

Third, if an emotional experience is evidence of things being a certain way or for the fact that things are that way, then it may still constitute a justificatory reason for action, independent of whether it does so for belief. With belief, we aim at the truth; with action, we aim at getting things right. What is true and what is right are different spheres of study and the requirements for what is a good reason to believe or to act can vary. For instance, both a true belief and a right action not only get things right in terms of believing truly or acting in light of real reasons, but can be justified or lack justification if they arise in the wrong kind of way. For belief, this may be something like not believing on the basis of strong evidence or believing as the result of bad reasoning. For action, this wrong kind of way may be something like not acting out of the right motivations or having a wrong conception of why the action is a good thing to do. So, how does an emotional experience as evidence fare with regards to action?

If an agent acts on the basis of an emotional experience, where the emotional experience provides her with access to facts because of its intentional nature and the way it implicitly represents those facts in its intentional content, then the emotional experience needs to be evidence for the fact. In Chapter 6, I introduced the example of John who backs away from the snake because he is afraid of it. I identified three potential facts that could be his reason. These were:

1. How things seem to him: There is a frightening snake (to me).

2. How things are, in terms of what response is merited: There is a frightening snake.

3. How things are, in terms of what relation holds: There is a dangerous(-to-me) snake.

For each of these options, we can ask two questions: Can an emotional experience be evidence for the reason in light of which the agent acts? If so, does it constitute a justificatory reason for the action? I will answer the first question for each potential reason, then present Brady's argument against supposing that the emotional experience constitutes a reason of any kind (for a judgement). I will question the principle

guiding Brady's argument and then illustrate how the emotional experience could constitute justificatory reasons of varying strengths for acting in light of the three reasons above.

Let us start with the first question of whether an emotional experience can be evidence for the reason in light of which the agent acts. With regards to the reason in 1, the question is whether an emotional experience can be evidence of how things seem to the agent. The answer is that it can indeed. It is very strong evidence of how things seem to the agent. If John experiences fear, then his experience is evidence of something being frightening to him. His fear is in fact constituted by finding something frightening to him.

The emotional experience is also evidence of the existence of the formal object in 2. This is because 'frightening' is response-dependent, and fear is the kind of response that is merited. If someone experiences fear, then it is the right kind of response to have to something frightening, and we expect someone to experience fear if there is something frightening. This does not yet show that the evidence is strong, however. Statistical correlation is exactly the kind of thing to expect if there is a causal link, but it remains weak evidence of a causal link. Nevertheless, if what is frightening is in part constituted by the agent having certain cares, and if having those cares also involves responding with fear when those cares are threatened, then a fear response would be fairly strong evidence for one element of the formal object, the fact that the agent is caring. A similar thing cannot be said for statistical correlation because causation is not in part constituted by statistical correlation.

It is less clear whether a fear response is evidence that an evaluative fact obtains.[14] Unlike being frightening, being dangerous(-to-me) does not call for a response of a certain kind and being dangerous is not in part constituted by the subject as caring (even though what cares the subject has may constitute what relation holds). However, being dangerous is still relative to the particular subject, either as an individual or as a member of a species. If experiencing fear is in part finding something dangerous, as I argued in Chapter 6, then an emotional experience of fear can be evidence for the fact that something is dangerous, relative to the subject and her cares. It just might not be strong evidence, just as statistical correlation is weak evidence for a causal link.

I have proposed that emotional experiences can be evidence of evaluative facts. In virtue of being evidence, they could constitute justificatory reasons for action on the basis of an emotional experience and in light of those evaluative facts.

[14]Elgin (2008) argues that an emotional experience *is* evidence for an evaluative fact like this, and she is one of Brady's main targets.

Brady, however, puts forward a second claim, which is that 'emotional experience doesn't, even in normal conditions, provide *any* kind of reason for evaluative judgement' (Brady 2013, 112). The argument in support of this claim is this. Emotional experiences, unlike perceptual experiences, are responses to reasons. As Brady notes, we think of recalcitrant emotions as 'irrational' because they are not responding to the reasons of which the agent is aware, and we think that certain reasons merit certain emotional responses. We do not think the same for perception. For instance, if I continue to see the stick in the water as bent even though I know it is straight, then I am not 'irrational' as I would be in the emotion case or, at least, I am not held to normative requirements of what my experience *should* be like.[15] Nor is there a sense in which seeing the stick as straight is somehow merited. Brady takes these considerations to illustrate the general point that:

> . . . the considerations that constitute good reasons for an emotional response are *equally* good reasons for the relevant evaluative judgement; and, by the same token, considerations that constitute good reasons for some evaluative judgement are equally good reasons for the relevant emotional response. (Brady 2013, 113)

For instance, the rhinkals raising its head to spit at John is both a good reason for him to feel fear and a good reason to judge that the snake is dangerous. But this, Brady argues, rules out that emotions can be reasons for evaluative judgements because 'such experience would appear to be capable of justifying itself' (Brady 2013, 113). That is, if feeling fear is a good reason to judge that the snake is dangerous, then it will also be a good reason to feel fear. The emotional experience ends up justifying itself, an undesirable commitment to be forced into. This is the 'bootstrapping' objection.

The principle guiding Brady's argument is that, if something is a good justificatory reason for one thing, then it will be a good justificatory reason for another. But as my arguments have aimed to show, this principle must be questioned. As I have already argued, the strength of the evidence and whether it constitutes a reason or a good or bad one will depend on what it is a reason for. Even accepting that an emotional experience does not constitute a good justificatory reason for itself, it does not follow that the same emotional experience does not constitute a good

[15] We do hold people to some requirements, however. When pointing out some object to someone who just cannot see it, we might find ourselves saying, 'You really should see it! Are you wearing your contact lenses?' Or look for other explanations for why the person cannot see it: are they hallucinating? Are they colour blind? Are they having a stroke? This suggests that we do have some requirements of what an experience should be like.

justificatory reason for something else, like a judgement or an action.

With this in mind, let us see how the three potential reasons, outlined in 1-3, fare and whether an emotional experience can constitute a justificatory reason for acting in the light of them, which was the second question above.

Starting with the first reason, which is the fact of how things seem to the agent, we see that an emotional experience does constitute a justificatory reason for acting on the basis of an emotion because of the way things seem to her, because the emotional experience is strong evidence of the way things seem to her. The way things seem to the agent is also a reason to judge that things seem a certain way to her. So, while in most contexts it may be a very bad kind of justificatory reason, it is still a reason and can at least explain—if not justify in a strong sense—the action as intelligible.

Does an emotional experience constitute a justificatory reason for acting on the basis of an emotion because of a fact about the formal object, the second potential reason? If being afraid entails that one experiences one's cares as being affected in a negative way, and experiencing fear is evidence for the fact that there is something affecting one negatively then, in the context of action, an emotional experience can constitute a justificatory reason for action because the experience draws attention to what cares are at stake which may require quick action. This does not mean that an emotional experience always constitutes a justificatory reason for action, however. It will depend on the context, not least on what the action is, just as statistical correlation may constitute a reason in favour of engaging in further study but not a reason in favour of releasing a press statement about a new health risk. In particular, finding something frightening may be a bad justificatory reason for acting if action is not immediately required. For instance, suppose you find frightening the prospect of a loved one's sudden death during an upcoming and risky operation. You call the local priest to organise last rites solely on the basis of experiencing fear. Such a case *does* call for further justification for your hasty and extreme action. But calling the priest on the basis of finding frightening the prospect of a loved one's death during an emergency life-reviving operation does *not* call for justification to the same extent, if at all: you needed to act quickly on a matter of deep importance to you, the import as indicated by your fear. But this is a matter of good and bad reasons, not a matter of the experience not constituting a reason at all.

Brady (2013, 114) does have something to say about the kind of response-dependent characterisation of the formal object which makes up the second reason, however. He argues that, if the characterisation of the formal object as response-dependent is correct, then an emotional experience cannot be a reason for a judge-

ment (setting aside action for now). This is because the emotional experience would be a reason to judge that something is the formal object, which is the same as saying that the emotional experience is merited, and we again run into the issue of the emotional experience justifying itself.

In order to respond to this argument, it is necessary to point out that, on a response-dependent picture, whether a certain response is merited can be assessed according to subject-independent features. Because of this, an emotional experience does not necessarily constitute a good justificatory reason for itself even though it could constitute a good reason for believing that an emotional response is merited. First, I argued in Chapter 6 that there are at least two important elements to our concept of the formal object. One element is that there is an evaluative relational property, like being dangerous, and the other element is that the agent actively cares, which affects how she finds things. This means that an emotional experience can be merited either because the agent's cares really are being affected or because the evaluative relation obtains. Simply because the agent experiences a situation in a certain way, however, does not entail that the situation is that way and that her response is merited. Second, with the particular sentimentalist line, there is also an agent-independent assessment of *when* a response is merited, which may be dependent on historical or cultural factors. Together, we see that, while the emotional experience is fairly strong evidence of the way the subject's cares are being affected and hence evidence of the formal object, it constitutes an easily defeasible reason to believe that the formal object obtains because there are other elements to the formal object of which the experience is not strong evidence, such as whether the response is appropriate relative to a context and culture or whether it is suitably grounded in an evaluative relation. Because of this, an emotional experience *can* constitute a justificatory reason for the belief about the formal object, but it does not follow that it then justifies itself in a problematic way because there are closely-connected defeating factors which determine whether the emotional experience is merited. It may ultimately be a bad reason, but it remains a reason nonetheless.

How does the third potential reason fare, the fact of what evaluative relation holds? I proposed that the emotional experience could be weak evidence for the fact. Just as statistical correlation is weak evidence for a causal link but does constitute a justificatory reason for believing that there is one, albeit not a good or sufficient one which silences the call for justification, an emotional experience could still constitute a justificatory reason for acting because something is dangerous, just not necessarily one which silences the call for justification. Further, if we question the guiding principle of Brady's bootstrapping objection as I have done, then even if

the emotional experience can constitute a justificatory reason for acting in light of that property or even believing that something has the property, the experience can be a bad justificatory reason for itself.

We therefore see that emotional experiences can be evidence of evaluative facts. In virtue of being evidence, they can constitute justificatory reasons. Whether they are good or bad reasons will depend on the context and what they are reasons for, but they can even sometimes be good justificatory reasons for action. Unlike the perceptual theorists, I do not require that the emotional experience should be sufficient reason or always a good reason, only that it can constitute a reason some of the time and, sometimes, even be a good reason.

Having discussed how an emotional experience can be suitably linked to the agent's reason for acting, I now turn to the second line of objection to our acting reasonably on the basis of an emotional experience.

7.2.2 Emotional experiences as providing prima facie entitlement to act

A different line of objection does not focus on emotions as reasons, but more on our entitlement to believe—or act—on their basis. Even if an emotional experience can be evidence of an evaluative fact, and even if an emotional experience can constitute a justificatory reason, it may still be the case that we are not entitled to act on the basis of an emotional experience.

This is the line pursued by Salmela (2011), who argues that we are not entitled to believe on the basis of an emotional experience in the same way as we are entitled to believe on the basis of a perceptual experience.[16] The core of his argument is this. The faculty of perception can get things wrong and so a perceptual experience is a defeasible reason for the perceptual belief. Nevertheless, we are entitled to believe on the basis of a perceptual experience because we can still specify conditions for when a perceptual experience will be defeasible as a reason or would not be evidence for the perceptual fact. In other words, we can establish when a perceptual experience gets things right and when it gets things wrong and, because we can do so, we are entitled to believe on the basis of the perceptual experience. The same is not true of emotional experiences, Salmela argues. We cannot give the conditions for when an emotional experience is accurate or inaccurate.[17] If we cannot specify

[16] Like Brady, his particular targets are the strong perceptual theorists who argue for an epistemic role for the emotional experience alone.

[17] As Salmela notes, not many have tried to give conditions or detailed conditions. Deonna (2006) stands out as giving the most comprehensive attempt, and it is thus with Deonna that Salmela is

what those conditions are, then we cannot establish when an emotional experience is evidence for an evaluative fact and we lack the *prima facie* entitlement to form a belief on the basis of the experience alone.

Salmela accepts the claim that perception gives us access to perceptual facts and emotions give us access to evaluative facts, and he accepts that both kinds of facts are relative to the agent's particular perspective. That perspective will determine what the conditions are and when an experience is defeasible as a reason. So why does he argue that we cannot give conditions for when an emotional experience is accurate or inaccurate? And how would this kind of objection apply to action?

The answer to the first question lies in how the perspectival nature of emotions and perception are different. The kind of perceptual fact that I become aware of through perception is perspectival because it is relative to my standpoint, in terms of my location and in terms of my physical constitution. We can specify what an agent's physical constitution and location are and how that affects the truth of the perceptual information they are receiving. Doing so is providing what we might call, following Deonna (2006), a 'frame of reference'. Using a frame of reference, we can identify what the defeaters for a perceptual experience as a reason will be, such as there being damage to the visual system, the mechanism being unreliable in a given context, or there being obstruction in the field of vision. Because we can identify conditions for when a perceptual experience will not track genuine information and is defeasible as a reason for assenting to the apparent information, we are entitled to rely on perceptual experiences in the cases where the conditions are in favour of the experience being correct. In terms familiar to my present discussion, we are entitled to believe on the basis of a perceptual experience because we can give the conditions for when the perceptual experience will be evidence for the fact and when it will not.

The relevant evaluative facts for emotion are also perspectival. They are relative to the agent's perspective, in terms of her location and to her capacities, but also in terms of her particular set of interests and background. However, the influence of the agent's cares and background on the evaluative facts make the emotional perspective substantially different to the perceptual perspective, and giving a frame of reference harder. Deonna has developed a frame of reference and it is this that Salmela challenges.

For Deonna, the frame of reference for an emotion's evaluative fact will be defined by species-relative facts as well as by individual-relative facts. For emo-

primarily concerned.

tion, unlike perception, the individual-relative facts will be dominant and include longstanding evaluative tendencies of the agent, such as being a racist or Christian, as well as her character traits or temperaments, such as being shy or irritable. If someone is a racist, for instance, then the Greek far-right party Golden Dawn's winning seats in the European Parliament presents a tangible and immanent good for her, and is exciting. If she is shy, then public speaking is a danger to her well-being, and is frightening. Defeaters for emotion will include false beliefs, phobias, traumas, etc.

Salmela argues, however, that our evaluative tendencies and character traits do not provide a frame of reference. This is because, in order to use them to establish an emotion's accuracy, we need to know if they *already* reflect evaluative facts about ourselves. Unlike facts about an agent's constitution and location, facts about evaluative tendencies and character require establishing that the agent herself endorses them: the defeaters for emotion 'go deeper in the psychology of the specific individual than the defeaters of perceptual experience' (Salmela 2011, 20). This requires what Salmela, following Deonna, calls a 'mastery of defeasibility' dependent on a degree of self-knowledge. As Salmela argues:

> My internalized character traits and longstanding values are obviously facts about (the present) me, but without an advanced 'mastery of defeasibility', these facts cannot provide a frame of reference for emotional responses that track evaluative facts for me. (Salmela 2011, 21)

If we cannot give a frame of reference for emotion, then we cannot identify when an emotional experience is evidence for an evaluative fact and when it is not; therefore, we are not entitled to believe on the basis of the experience alone.

Salmela's focus is on belief but he raises important worries for action. If we cannot specify under what conditions an emotional experience is evidence, regardless of whether it constitutes a reason when it is evidence, we are not entitled to act on the basis of the emotion alone. Nevertheless, it is not the case that we need a frame of reference as stringent as Salmela proposes for belief and this makes the worries less forceful.

To see this, let us go through how the reasons 1-3 I identified above fare, and see whether we are entitled to act in light of them on the basis of an emotional experience. We can set aside cases of the first kind of reason at the outset. These are facts about how something seems to the agent. We can set them aside because they are uninteresting and we can specify the conditions for their being evidence: the agent experiences the situation as being the way she experiences it.

Turning to the other two potential reasons, I shall argue that, contrary to what Salmela suggests, we can in fact provide enough of a framework to generate a *prima facie* entitlement. Salmela's target is the dimension of assessment of an emotion as accurate or inaccurate and I shall argue two points. First, even if we cannot give comprehensive conditions for accuracy of an emotion, we can give some conditions to develop a frame of reference and, second, there are other ways an emotional experience can fail to give entitlement. Taken together, we have a robust picture of when we can be entitled to act on the basis of an emotional experience.

Starting with the first point, which is that we can still give some conditions for accuracy and a frame of reference. The role of the frame of reference is to identify conditions which will affect the accuracy of the experience; its role is not of something to be used *before* one is entitled to do anything on the basis of the experience, or something to consult *before* going ahead with a belief or an action. This means that we can give some basic conditions for the accuracy of an emotional experience which allow us to specify when an experience is not accurate even if the agent is unaware at the time whether those conditions hold or not. Doing so is sufficient to generate entitlement. For instance, Deonna proposes using other people's judgements about ourselves as a way of establishing whether we are acting or believing in a way that is consistent with our evaluative tendencies and character traits. If someone who sees herself as shy is excited by the prospect of giving an address to a large crowd, it may be the case that her excitement is getting things wrong. However, if her peers judge her not to be shy, then her excitement may in fact be accurate and, indeed, she may learn something about herself through her experience. We can also establish what those evaluative tendencies and character traits are by giving counterfactual or hypothetical conditions, as Jones (2003) does. She argues that it is sufficient that:

> ...the agent's disposition to reflective self-monitoring are such that she would not rely on that first order sub-system [such as emotion or perception] were it reasonable for her to believe that it failed to reason-track. That is, her conscious reflective capacities exert *regulative guidance* over the first-order mechanisms, stepping in when necessary to discount those mechanisms, and where possible, to recalibrate them into reason-tracking mechanisms through habituation. Sometimes this guidance may remain 'virtual'—that is, revealed in how the agent would behave in various counterfactual circumstances (were she to have evidence that they are unreliable, for example, but evidence she never gets since they are reliable). (Jones 2003, 195-96)

We therefore *do* have ways to establish what evaluative facts are true of the subject, such as through other people's judgements about us or our own reflective self-monitoring. The subject need not be aware of the conditions at the time of believing or acting and may not herself have mastery of defeasibility, but she still has ways to assess her experience for accuracy. In this way, we can give less stringent conditions for when an emotional experience does accurately provide access to facts about the formal object or the evaluative relational property, the second and third potential reasons.

Even if these ways of establishing what facts are true of the subject are not sufficient to generate entitlement based on whether an emotion is accurate or not, there are other ways in which an emotional experience can fail to give entitlement to act. This is the second point.

The first way, in addition to establishing accuracy conditions, is to assess an emotion for appropriateness, which was one of the dimensions of assessment I discussed in Chapter 6 and applies especially to the second potential reason, a fact about the formal object characterised as response-dependent. If the formal object is understood as response-dependent, then some of our conditions for accuracy will be conditions regarding when such a response is merited. This will include conditions for whether the evaluative relation which makes up one element of the concept of the formal object obtains; conditions for when the emotion is elicited in the right kind of way, such as through the subject actively caring rather then responding to a chemical injection; and conditions about the social and cultural context the agent is in. In particular, the last we can establish without a mastery of defeasibility because the conditions for the appropriateness of an emotion are not 'deep within the psychology of the individual'. If the racist experiences fear when she encounters someone of a different ethnicity, for instance, regardless of whether the person really is a threat to the racist's particular commitments and whether the racist finds her frightening, she is not entitled to believe or act because the person does not merit fear and is not actually frightening.

This level of assessment is open to emotions because, unlike perception, emotions are responses to reasons and do have a normative element, as Brady points out and I mentioned in section 7.2.1. Our entitlement to believe or act on the basis of an emotional experience is thus unlike the entitlement for perception because it is not limited to whether the experience is accurate or not. It is also dependent on whether the emotion is one the agent ought to have. And this is true even if, for example, the racist who acts on his inappropriate fear ultimately does good in terms of what the action brings about, because his motivation will be wrong. Someone who is not

prone to reflective self-monitoring may also have less entitlement to act on the basis of her emotional experience because it is inappropriate on an epistemic level for her to do so.

Further, we can assess an emotion for intelligibility with regards to both the second and third potential reasons, facts about the formal object and about the evaluative relation. If the emotional experience is not intelligible, then it may be defeasible. For instance, if the emotional experience is in conflict with other beliefs that the agent holds, then one of those mental states is not providing her with a reason, or they all provide reasons but one of them will be defeasible. For instance, Clare the agoraphobe experiences fear when in open spaces but also believes that the situation is not frightening and that it is not dangerous. If she is to be subjectively rational, she needs to weigh up her considerations in order to assess which are and are not reasons or good reasons, and thus she is not entitled to act in light of the reason that the situation is frightening or that it is dangerous on the basis of the emotional experience *alone*. However, as discussed in section 7.2.1, we will sometimes need to act promptly and in such cases a lack of conflict is sufficient to generate an entitlement to act on the basis of the emotional experience.[18] The kind of entitlement is independent of the emotion's accuracy and the frame of reference.

We therefore see that intelligibility and appropriateness set up some basic conditions for entitlement and that we can give some general conditions for a frame of reference for both the second and third potential reasons. The kind of entitlement to act on the basis of an emotional experience is not the same as the kind of entitlement for perceptual experience, but there is entitlement nonetheless.

In summary, I have used my three potential reasons to assess the objections that emotional experiences do not provide evidence for an evaluative fact or constitute sufficient justificatory reasons for an action, and the objection that we are not entitled to act on the basis of an emotional experience alone. We have seen that each of the potential reasons has its own nuances, strengths and weaknesses, but that it is nevertheless the case that an emotional experience can be evidence and can constitute a justificatory reason, and that we can have *prima facie* entitlement to act on the basis of an emotional experience.

[18]Patricia Greenspan (2011) in fact argues that the emotional experience provides a critical (or normative) reason to act and has an important role in helping us to act timeously.

7.3 Conclusion

I have argued that the intentional nature of emotions gives rise to three potential reasons for which the agent could act. All three reasons are evaluative facts, although the first one, concerning how things appear to the agent, is ultimately uninteresting. It is not usually the reason an agent would cite to explain her action except perhaps as a last resort to show how her action was intelligible. Because of this, I have not given it much attention and shall now set it aside completely.

I have argued that, by being committed to the way things appear through caring and by being disposed to believe the putative fact that an emotion presents the agent with, we can act in light of that fact, which is the reason, when experiencing an emotion without first forming a belief about the evaluative fact. Lastly, I have argued that emotional experiences can be evidence for the evaluative facts, can constitute justificatory reasons for acting on those facts, and that we are *prima facie* entitled to do so.

What is left is to return to the examples of Category E where the emotion provides the agent's reason, but not as a state to be addressed. The first two remaining examples involve love:

- *Richard bought flowers for Katherine because he loved her.*
- *Natasha pulls her life together out of love for her mother.*

While the love in the first example could be an attitude, I shall interpret it as a moment of occurrent love: Richard thinks of Katherine with a surge of fondness and buys her flowers. He acts on the basis of an experience of love.

What would be his reason? The formal object of love could be cashed out as being affection-inspiring, and the evaluative relation would be something like being value-giving or being a source of great value in one's life. So, Richard buys flowers for Katherine because she is affection-inspiring or because she is the source of great value in his life. By feeling love, he is committed to her being of value to him and he acts on the basis of that commitment.

His experience of love is evidence for the fact that she is affection-inspiring, because one of its elements is the subject actively caring and because it merits that kind of response. It is evidence for the fact that she is of great value to him because part of loving is valuing. Because it is evidence, it does constitute a justificatory reason for his action. Is it a good or a bad justificatory reason? On one hand, it is not conclusive evidence. But on the other hand, it could still be sufficient

evidence. As mentioned in section 7.2.1, it is striking that positive emotions tend not to have the same call for justification as negative emotions: the explanation that Richard bought flowers for Katherine because he loves her does not call for further justification. In fact, if one were to say, but *why* does he love her? and push for further reasons, I would think that the person inquiring was being particularly obtuse or I would become curious about her history with Katherine, suspecting that there was some bad blood between them which makes her think that Katherine is particularly unloveable. Finally, Richard is *prima facie* entitled to act on the basis of his experience of love. Indeed, if he were to form other beliefs about Katherine's value and what he should do, we might even think that his action is poorer for not being a direct expression of his love. His motivation would be wrong.

This first example helps illustrate the way in which an attitude can also provide access to an agent's reasons, as with Natasha's love for her mother in the second example. Natasha is distraught over the death of Prince Andrey. She becomes ill and loses all purpose to her life. However, her love for her mother 'unexpectedly showed her that the essence of life—love—was still active within her' (Tolstoy 2001, 933). Through her love for her mother, Natasha starts to pull her life together. She is neither avoiding or encouraging the state, nor expressing it. Yet, it is true to say that she acts as she does *because she loves her mother*, where the love explains more than that she has a particular attitude.

In both of these examples, the agent's reason will be the loved one as affection-inspiring or the loved one as a source of great value. Even if Natasha does not experience an occurrent state of love as she pulls her life together, she nevertheless is committed to the value that her mother brings her, as is seen in her other attitudes and beliefs, as well as her desire to be better, perhaps in order to be worthy of her mother's love or more capable herself of participating in an affectionate relationship. She therefore does act in light of either the fact that her mother is affection-inspiring or the fact that her mother is value-giving.

Further, her attitude towards her mother is evidence, not necessarily sufficient evidence, for the fact that her mother is loveable and value-giving in her life; if her mother were not that way, then we would expect her attitude to be different.

Again, as a reason to pull her life together, while not an experience of love, her attitude of love does constitute a justificatory reason for why she does what she does. Indeed, like Richard's love, acting on the basis of love is acting out of worthy motivations. We can see this in a kind of converse case where someone's love is not a strong enough reason to carry on with things. A common criticism of a suicide is that the act is selfish because it leaves behind the person's friends

and family who love her. The underlying thought to such a criticism seems to be something like, if she really loved her friends and family, then she would not have done what she did. But if this is a criticism of such action, then presumably carrying on despite one's own debilitating condition but on the basis of one's love for others is a sufficient justificatory reason for doing so. Even rejecting that an action like suicide is selfish, the case still brings out how love for one's friends and family can constitute a justificatory reason for acting in light of the reason that there is value in one's life which, in the case of suicide, is either overlooked or not strong enough to count in favour of carrying on. Cases where it seems that the agent ought to act on the basis of love for others, or where her not doing so is somehow distressing, illustrate that sometimes, like Natasha, one is *prima facie* entitled to act on the basis of love.

The next example is closely related to the love examples:

- *Huck Finn does not turn Jim over to the authorities because he feels compassion for Jim.*

Huck Finn does not form other beliefs about Jim or what he ought to do, he simply acts out of his compassion. His reason would be grounded in the relation that Jim stands to him, as Jim needing kindness that Huck can give. He is committed to the plight of Jim through caring, and could offer it as a reason for why he goes against what he thinks he should do, even while he happens to think that it is a bad reason.

Feeling compassion is evidence that Jim needs kindness, because Jim's situation is such that it generates a merited response. Again, the evidence may not constitute a sufficient justificatory reason depending on the context and what it is a reason for. Because someone who needs compassion may often need it immediately, the experience could easily be a good justificatory reason for helping that person. And, like with the examples of love, acting on the basis of compassion rather than prudence or expedience is acting with the right kind of motivation and he is *prima facie* entitled to act on the basis of his compassion. Unlike love, there would potentially be more defeaters to the entitlement because compassion is not just about someone who is valuable, but about someone who needs kindness. It is more complex, allowing more scope for error.

The next example involves a more negative emotion:

- *Alice challenges her boss because she is morally indignant at the injustice.*

Alice challenges her boss because she feels moral indignation about the injustice he has committed by not giving her colleague a deserved payrise. Here, her reason

is that there has been an injustice, understood as meriting a moral response, or an unfairness. She is committed to the way there has been an injustice or an unfairness through caring, and can offer it as a reason for why she challenges her boss.

So long as her moral compass is suitably tuned, her indignation is evidence that the boss's action was unjust. Unlike the cases of love and compassion, however, her moral indignation, alone, is a bad justificatory reason for her action of challenging her boss. An emotion like moral indignation is cognitively complex and dependent on a moral code that will be independent of Alice's more direct cares and interests. Her direct care would be to do what is right, but what is required in order to do what is right goes beyond her and is contestable. Her emotional experience thus would have many defeaters, and defeaters that do not apply directly to her emotion but directly to her system of values. Nevertheless, so long as her emotion is not in conflict with her other beliefs, it is in line with her evaluative tendencies and character traits as well as with what is just and unjust, and it is appropriate given that an injustice has happened, she is entitled to act on the basis of her emotion.

Finally, the last example:

- *Emily quits the PhD programme because she is sad and ill-motivated.*

Emily, a PhD student in chemistry, finds herself feeling sad and ill-motivated in her endeavours, despite having always believed that she should pursue a PhD. She sees the feelings that make her want to leave the programme as groundless even though, unknown to her, they are triggered by factors that would be good reasons not to be doing the PhD. She eventually leaves the programme on an emotional impulse and judges herself to be acting irrationally. When she looks back years later, however, she views her persistence as irrational and takes as her reasons for acting the reasons behind her emotional state (Arpaly 2000, 504). Emily seems to act akratically. Nevertheless, she acts in light of a reason, which was that the PhD programme was creating disvalue in her life. In the example as originally described, she does not originally think that she has acted for a reason. However, the reason dawns on her over time (Arpaly 2000), which can be interpreted as indicating that she was aware of her reason but on a subconscious level. Assuming that such an interpretation of the example holds up, it is not a case that she constructs a reason to explain her action and gives a post-hoc rationalisation. She genuinely did act in light of a reason.

Her sadness is also evidence of the disvalue the PhD creates in her life, although it is evidence that she fails to appreciate or which she ignores. Her emotional experience, because it is evidence of the disvalue, does constitute a justificatory reason.

However, she is not entitled to act on the experience alone because it is in conflict with her other beliefs and mental states.

Working through the examples in this way, we see the variety of ways in which an emotion can provide access to a reason, how an agent can act in light of a reason when acting on the basis of an emotional experience, how the emotional experience can be evidence for an evaluative fact, how it can constitute a justificatory reason, and how we can be entitled to act on the basis of the experience. With these different ways of assessing an action on the basis of an emotion, such actions cannot be dismissed out of hand as not fully rational. Moreover, an explanation in terms of an emotion can be an explanation alluding to the agent's motivating reason, and in that way, an explanation which shows how the agent was acting rationally and reasonably.

CONCLUSION

In the Introduction, I set up my project in this thesis as bridging part of the gap between our understanding of the functional role of emotions and our conception of rational action. In particular, I aimed at showing how, if emotions function in a way whereby they provide us with information about how things are, we can act rationally when acting on the basis of them. By arguing that acting rationally is acting in light of a reason and by focusing on explanations of action, I have argued that we can act in light of a reason when acting on the basis of an emotion: a reason provided by an emotion in virtue of its intentional nature. I divided my argument into three parts.

In Part I, I argued that there is conceptual space for emotions to have a reason-giving role in action and in explanations of action. In Chapter 1, I developed the claim that our concept of emotion is of something rich and diverse. In Chapter 2, I introduced five categories of different ways in which we explain action by citing an emotion. Taking these two claims together, we saw that, even if an emotion plays one kind of role in one kind of explanation, it need not play the same role in all explanations. In Chapter 3, I illustrated how, despite the variety of potential roles, emotions have traditionally been relegated to only one or perhaps two roles. I argued that limiting the roles in this way cannot accommodate the differences between the categories of examples of explanation of Chapter 2 and that, if we properly embrace the intentional nature of emotions, the conceptual space for acting for a reason when acting on the basis of an emotion is opened up.

In Part II, I covered important background commitments regarding what emotions are, where I argued in Chapter 4 that we should adopt a weak cognitive account of emotions such as a perceptual account, and regarding rational action, where I argued in Chapter 5 that acting for a reason is acting in light of a fact. I concluded Part II by describing more formally what the differences between the categories of explanation are. I stated more precisely how emotions could plausibly provide access to the agent's motivating reason because of their intentional nature, as well as how an explanation citing an emotion could be an explanation of the action as

reasonable.

In Part III, I developed the details of my view. In Chapter 6 I examined in detail the concept of the formal object of an emotion in order to establish what reasons an emotion could provide in virtue of its intentional nature, and showed how these reasons were facts in light of which we could act. Finally, in Chapter 7, I argued that we could act reasonably when acting solely on the basis of an emotional experience and that we could act in light of a reason when acting solely on the basis of an emotion.

There are several further avenues of research which my arguments and claims in this thesis open up or to which they could contribute. I shall close by mentioning five of them.

The first avenue of research is to explore a potential objection to my argument that we can act in light of a reason when acting from an emotion. This is an objection based on the idea that, in order to act for a reason, we must have knowledge of that reason (see, for example, Hyman 1999; Hornsby 2007). It is a potential objection because I have said nothing about whether we know the reasons emotions provide us with, and it may well turn out that we do not know those reasons.[19] If so, then we could not act in light of a reason when acting on the basis of an emotion. It is not currently clear to me to what extent we would need knowledge in order to act for a reason, and whether we could have knowledge of the reasons even if, say, emotional experiences do not provide good reasons for belief. In order to explore this objection, I would need to examine in detail the arguments that we must have knowledge of a reason in order to act for it, and see how my proposal is affected.

The second avenue of research is to extend my focus on emotions beyond just occurrent emotion states. While my focus in this thesis has been necessarily narrow, some of my claims and arguments could shed light on the relation between action and emotional attitudes, action and moods, and even action and emotionally-based character traits.

For example, in Chapter 3 I argue that acting out of an emotion as a motive-proper is acting with a certain spirit. This could be drawn on to examine character traits in more detail and the proposal that explaining an action as being done out of a particular character trait reflects a 'spirit' with which we act. Or, if emotions involve a commitment to the way things appear, as I argue, then an emotional attitude would also involve a commitment to how the things one cares about are being affected. This could be put to use in other areas of philosophy, such examining more closely

[19]The case of Emily the PhD student whose reasons 'dawn on her', for instance, does not clearly know the fact that the PhD is bad for her.

how and what reasons interpersonal relationships, such as ones of love, provide us with.[20]

The third avenue of research is closely related to this latter idea. I have examined emotions as they relate to the agent's motivating reason, but a more thorough investigation of emotions as they relate to normative reasons is needed. Could an emotion provide a normative reason for action? Patricia Greenspan (2011) argues that emotions provide what she calls critical reasons for action in virtue of their discomfort. I do not think that her arguments are successful because she fails to distinguish satisfactorily between the cases in Category D of my examples, cases where we act to address the physical manifestation of an emotion and cases where we act to address an emotion as an intentional state.[21] I would like to examine more closely whether there are times when an emotion, in itself, provides a normative reason to act. Perhaps an attitude of love, for example, provides normative reasons to protect loving relationships. This will feed into whether we are indeed entitled to act on the basis of an emotional experience, as discussed in Chapter 7.

The fourth avenue of research is to explore the consequences of my view of the concept of the formal object as a response-dependent concept with one of the elements an evaluative relational property and another element the subject as actively caring. For example, in their recent book on emotions, Julien Deonna and Fabrice Teroni (2012) argue that emotions are not about their formal object and argue in favour of a view where emotions are a kind of attitude. It is in virtue of the kind of attitude that an emotion is, they argue, that the formal object provides the correctness conditions for the experience. It is not because emotions are *about* the formal object, where the emotional content is what has correctness conditions.

The kind of view I am proposing is, on the one hand, in conflict with what Deonna and Teroni propose because I allow that emotions can be about evaluative properties; on the other hand, it centres around a similar idea to theirs, namely the idea that the formal object is a property of the emotional experience rather than what the emotion explicitly represents. One further avenue of research which I would like to pursue is a more detailed examination of Deonna and Teroni's view in order to see how my understanding of the formal object fits in. Nevertheless, their view is a recent move-away from the popular perceptual accounts. I am attracted to it superficially, and believe that my examination of the formal object could help

[20]Thank you to Sandy Koll of Johns Hopkins University for discussing her work on reasons and relationships with me and making me think about this potential avenue.

[21]I have argued against her view in papers presented at two conferences, Open Minds VII at the University of Manchester in June 2012 and the British Postgraduate Philosophy Society annual conference in September 2012.

adjudicate the conflicting accounts of emotions.

The fifth avenue of research is to focus more exclusively on explanations. In Chapter 5, when discussing expressive action and the example of the man who rolls around in his dead wife's clothing out of grief, I argued that the man's emotion, grief, provides a desirability characterisation of his action. And in Chapter 3 when discussing objections to belief-desire accounts of emotions, I quoted Andrea Scarantino:

> . . . being motivated to get away from a snake while fearing it is inseparable from representing it as dangerous and not freely employable in practical inferences, as the desire to get away from a snake would be. (Scarantino 2010, 361)

Taking these two ideas, a proposal I would like to explore in more depth is the proposal that an explanation citing an emotion is fundamentally different from an explanation citing a belief because the emotion, but not the belief, provides both the agent's reason in favour of acting, as well as the desirability characterisation for why she saw those considerations and action in that way. Very loosely put, the agent runs away because *the snake is frightening*, but she sees running away as a good thing to do because *the snake is frightening*—not because of some separate desire, such as a desire to get away. If this idea could be developed, then an emotion explanation would be a very distinct kind of action explanation.[22]

These avenues of research, together with the other contributions I identified in the Introduction, illustrate the significance of this thesis project. But perhaps the most important contribution is that it acknowledges our emotional nature and contributes towards a robust conception of rational agency which is able to incorporate emotions into an account of how we do act in many of the cases when we act rationally.

[22] I begin to explore this idea in a paper presented at the Georgetown University Philosophy Conference on Reasons and Reasoning in April 2013 and the Berkeley-London Graduate Conference in May 2013, where I explore an analogy with John McDowell's (1979) arguments about virtue as the virtuous agent's reason for acting.

REFERENCES

Alvarez, M. (2008). 'Motivating Reasons and the Ambiguity of Belief'. *Philosophical Explorations 11*, 53–65.

Alvarez, M. (2009). 'Acting Intentionally and Acting for a Reason'. *Inquiry 52*, 293–305.

Alvarez, M. (2010). *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford: Oxford University Press.

Anscombe, E. (1957). *Intention* (2nd ed.). Oxford: Oxford Basil Blackwell.

Aquinas, T. (1967). *Summa Theologiæ: A Concise Translation*, Volume XIX. London: Blackfriars.

Aristotle (1983). 'On the Soul'. In C. Calhoun and R. Solomon (Eds.), *What Is An Emotion? Classic Readings in Philosophical Psychology*, pp. 48–50. Oxford: Oxford University Press.

Aristotle (2007). *On Rhetoric*. Oxford: Oxford University Press.

Arnauld, A. and P. Nicole (1996). *Logic or the Art of Thinking*. Cambridge: Cambridge University Press.

Arpaly, N. (2000). 'On Acting Rationally Against One's Best Judgment'. *Ethics 110*, 488–513.

Audi, R. (2002). *The Architecture of Reason: The Structure and Substance of Rationality*. Oxford: Oxford University Press.

Bacon, F. (1893). *The Advancement of Learning*. London: Cassell and Company.

Barrett, L. F., B. Mesquita, K. N. Ochsner, and J. J. Gross (2007). 'The Experience of Emotion'. *The Annual Review of Psychology 58*, 373–403.

Betzler, M. (2009). 'Making Sense of Actions Expressing Emotions'. *Dialectica 61*, 447–66.

Brady, M. (2013). *Emotional Insight: The Epistemic Role of Emotional Experience*. Oxford: Oxford University Press.

Broome, J. (1999). 'Normative Requirements'. *Ratio 12*, 398–419.

Broome, J. (2005). 'Does Rationality Give Us Reasons?'. *Philosophical Issues 15*, 321–37.

Byrne, A. (2001). 'Intentionalism Defended'. *Philosophical Review 110*, 199–240.

Cannon, W. (1927). 'The James-Lange Theory of Emotions: A critical examination and an alternative theory'. *The American Journal of Psychology 39*, 106–24.

Crane, T. (1998). 'Intentionality as the Mark of the Mental'. In A. O'Hear (Ed.), *Current Issues in Philosophy of Mind*, pp. 229–51. Oxford: Oxford University Press.

Crane, T. (2009). 'Is Perception a Propositional Attitude?'. *Philosophical Quarterly 59*, 452–69.

Crane, T. (2014). *Aspects of Psychologism*. Cambridge, MA: Harvard University Press.

Damasio, A. (1994). *Descartes' Error*. London: Vintage Books.

Dancy, J. (2000). *Practical Reality*. Oxford: Oxford University Press.

Dancy, J. (2004). *Ethics Without Principles*. Oxford: Oxford University Press.

D'Arms, J. (2005). 'Two Arguments for Sentimentalism'. *Philosophical Issues 15*, 1–21.

D'Arms, J. and D. Jacobson (2000). 'The Moralistic Fallacy: On the 'appropriateness' of emotions'. *Philosophy and Phenomenological Research 61*, 65–90.

Davidson, D. (1963). 'Actions, Reasons and Causes'. *Journal of Philosophy 60*, 685–700.

Davidson, D. (2004). *Problems of Rationality*. Oxford: Oxford University Press.

de Sousa, R. (1987). *The Rationality of Emotion*. The MIT Press: Cambridge, Massachussetts.

de Sousa, R. (2014). 'Emotion'. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy (Spring 2014 Edition)*. Accessed 26 February 2014, http://plato.stanford.edu/archives/spr2014/entries/emotion/.

Deigh, J. (2010). 'Concept of Emotions in Philosophy and Psychology'. In P. Goldie (Ed.), *The Oxford Handbook of Philosophy of Emotion*, pp. 17–40. Oxford: Oxford University Press.

Dennett, D. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

Deonna, J. (2006). 'Emotion, Perception and Perspective'. *Dialectica 60*, 29–46.

Deonna, J. and F. Teroni (2012). *The Emotions: A Philosophical Introduction*. London: Routledge.

Descartes, R. (1985). 'The Passions of the Soul'. In *Philosophical Writings of Descartes Volume 1*. Cambridge: Cambridge University Press.

Dickens, C. (1905). *Hard Times*. London: Chapman & Hall.

Döring, S. (2003). 'Explaining Action by Emotion'. *Philosophical Quarterly 53*, 214–30.

Döring, S. (2007). 'Seeing What to Do: Affective perception and rational motivation'. *Dialectica*, 363–94.

Döring, S. (2008). 'Conflict Without Contradiction'. In G. Brun, U. Doğuoğlu, and D. Kuenzle (Eds.), *Epistemology and Emotions*, pp. 83–103. Aldershot: Ashgate.

Döring, S. (2010). 'Why Be Emotional?'. In P. Goldie (Ed.), *The Oxford Handbook of Philosophy of Emotion*, pp. 283–302. Oxford: Oxford Basil Blackwell.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge: MIT Press.

Dretske, F. (1986). 'Misrepresentation'. In R. Bogdan (Ed.), *Belief: Form, Content, and Function*, pp. 17–34. Oxford: Clarendon Press.

Ekman, P. (1977). 'Biological and Cultural Contributions to Body and Facial Movement'. In J. Blacking (Ed.), *The Anthropology of the Body*, pp. 39–84. London: Academic Press.

Ekman, P. (1992). 'An Argument for Basic Emotions'. *Emotion and Cognition 6*, 169–200.

Elgin, C. Z. (2008). 'Emotion and Understanding'. In G. Brun, U. Doğuoğlu, and D. Kuenzle (Eds.), *Epistemology and Emotions*, pp. 33–50. Aldershot: Ashgate.

Evans, J. S. and K. E. Stanovich (2013). 'Dual-Process Theories of Higher Cognition: Advancing the debate'. *Perspectives on Psychological Science 8*, 223–41.

Everson, S. (2010). 'Motivating Reasons'. In T. O'Connor and C. Sandis (Eds.), *A Companion to the Philosophy of Action*, pp. 145–52. Chichester: Blackwell Publishing.

Fodor, J. (1984). 'Semantics: Wisconsin style'. *Synthese 59*, 231–50.

Fodor, J. (1989). *Psychosemantics*. Massachussetts: The MIT Press.

Frijda, N. (1986). *The Emotions*. Cambridge: Cambridge University Press.

Gendler, T. (2008). 'Alief and Belief'. *The Journal of Philosophy 105*, 634–63.

Goldie, P. (2000). *The Emotions: A Philosophical Investigation*. Oxford: Oxford University Press.

Green, O. H. (1992). *The Emotions: A Philosophical Theory*. Dordrecht: Kluwer Academic Publishers.

Greene, J. (2007). 'The Secret Joke of Kant's Soul'. In W. Sinnott-Armstrong (Ed.), *Moral Psychology 3: The Neuroscience of Morality. Emotion, Brain Disorders and Development*. Boston: MIT Press.

Greenspan, P. (1988). *Emotions and Reasons: An Inquiry into Emotional Justification*. New York: Routledge.

Greenspan, P. (2011). 'Craving the Right: Emotions and moral reasons'. In C. Bagnoli (Ed.), *Morality and the Emotions*, pp. 39–61. Oxford: Oxford University Press.

Griffiths, P. (1997). *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.

Hacker, P. (2004). 'The Conceptual Framework for the Investigation of Emotions'. *International Review of Psychiatry 16*, 199–208.

Haidt, J. (2001). 'The Emotional Dog and its Rational Tail'. *Psychological Review 108*, 814–34.

Haukioja, J. (2013). 'Different Notions of Response-Dependence'. In M. Hoeltje, B. Schnieder, and A. Steinberg (Eds.), *Varieties of Dependence—Ontological Dependence, Grounding, Supervenience, Response-Dependence*, pp. 167–90. Munich: Philosophia Verlag.

Helm, B. (2001). *Emotional Reason: Deliberation, Motivation, and the Nature of Value*. Cambridge: Cambridge University Press.

Helm, B. (2009). 'Emotions as Evaluative Feelings'. *Emotion Review 1*, 248–55.

Hornsby, J. (2007). 'Knowledge in Action'. In A. Leist (Ed.), *Action in Context*, pp. 285–302. Berlin: De Gruyter.

Hume, D. (2006). *A Treatise of Human Nature*. Oxford: Oxford University Press.

Hursthouse, R. (1991). 'Arational Actions'. *The Journal of Philosophy 88*, 56–68.

Hyman, J. (1999). 'How Knowledge Works'. *Philosophical Quarterly 50*, 433–51.

Izard, C. (2010). 'The Many Meanings/Aspects of Emotion: Definitions, functions, activation, and regulation'. *Emotion Review 2*, 363–70.

James, S. (1997). *Passions and Action: The Emotions in Seventeenth-Century Philosophy*. Oxford: Clarendon Press.

James, W. (1884). 'What is an Emotion?'. *Mind 9*, 188–205.

Johnston, M. (1989). 'Dispositional Theories of Value'. *Proceedings of the Aristotelian Society Supplementary Volume 63*, 139–74.

Jones, K. (2003). 'Emotion, Weakness of Will, and the Normative Conception of Agency'. In A. Hatzimoysis (Ed.), *Philosophy and the Emotions*, pp. 181–200. Cambridge: Cambridge University Press.

Jones, K. (2008). 'Quick and Smart? Modularity and the Pro-Emotion Consensus'. *Canadian Journal of Philosophy Supplementary Volume 32*, 3–27.

Jost, J. T., L. A. Rudman, I. V. Blair, D. R. Carney, N. Dasgupta, J. Glaser, and C. D. Hardin (2009). 'The Existence of Implicit Bias is Beyond Reasonable Doubt: A refutation of ideological and methodological objections and executive summary of ten studies that no manager should ignore'. *Research in Organizational Behavior 29*, 39–69.

Kahnemann, D. (2011). *Thiking, Fast and Slow*. London: Penguin.

Kenny, A. (1976). *Action, Emotion and Will*. London: Routledge & Kegan Paul.

Kenny, A. (1989). *The Metaphysics of Mind*. Oxford: Clarendon Press.

Kolodny, N. (2005). 'Why Be Rational?'. *Mind 114*, 509–63.

Lazarus, R. (1991). *Emotion and Adaptation*. Oxford: Oxford University Press.

Lazarus, R. (2001). 'Relational Meaning and Discrete Emotions'. In K. Scherer, A. Schorr, and T. Johnstone (Eds.), *Appraisal Processes in Emotion: Theory, Methods, Research*, pp. 37–67. Oxford: Oxford University Press.

Leventhal, H. and K. Scherer (1987). 'The Relationship of Emotion to Cognition: A functional approach to a semantic controversy'. *Cognition and Emotion 1*, 3–28.

Lewis, M. (2010). 'The Emergence of Human Emotions'. In M. Lewis, J. M. Haviland-Jones, and L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed.)., pp. 304–19. New York: The Guilford Press.

Linquist, S. and J. Bartol (2013). 'Two Myths About Somatic Markers'. *British Journal for the Philosophy of Science 64*, 455–84.

Littlejohn, C. (2012). *Justification and the Truth-Connection*. Cambridge: Cambridge University Press.

Lyons, W. (1980). *Emotion*. Cambridge: Cambridge University Press.

Marks, J. (1982). 'A Theory of Emotion'. *Philosophical Studies 42*, 227–42.

McDowell, J. (1979). 'Virtue and Reason'. *Monist 62*, 331–50.

McDowell, J. (1984). 'Values and Secondary Qualities'. In T. Honderich (Ed.), *Morality and Objectivity*. London: Routledge.

Millikan, R. G. (1989). 'Biosemantics'. *The Journal of Philosophy 86*, 281–97.

Moors, A. (2014). 'Flavors of Appraisal Theories of Emotion'. *Emotion Review 6*, 303–307.

Mulligan, K. (2010). 'Emotions and Values'. In P. Goldie (Ed.), *The Oxford Handbook of Philosophy of Emotion*, pp. 475–500. Oxford: Oxford University Press.

NHS (2011). Phobias - Claire's story. Accessed 17 December 2013, www.nhs.uk/Conditions/Phobias/Pages/Clairesstory.aspx.

Nussbaum, M. (2001). *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.

Oatley, K. and P. N. Johnson-Laird (1987). 'Towards a Cognitive Theory of Emotions'. *Cognition and Emotion 1*, 29–50.

Ortony, A. and T. Turner (1990). 'What's Basic About Basic Emotions?'. *Psychological Review 97*, 315–31.

Osman, M. (2013). 'A Case Study: Dual-Process Theories of Higher Cognition—Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science 8*, 248–52.

Parfit, D. (1997). 'Reasons and Motivation'. *Proceedings of the Aristotelian Society Supplementary Volume 71*, 99–130.

Pettit, P. (1991). 'Realism and Response-Dependence'. *Mind 100*, 587–626.

Plutchik, R. (2001). 'The Nature of Emotions'. *American Scientist 89*, 344–50.

Prinz, J. (2000). 'Duality of Content'. *Philosophical Studies 100*, 1–34.

Prinz, J. (2002). *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge: MIT Press.

Prinz, J. (2004). *Gut Reactions*. Oxford: Oxford University Press.

Raz, J. (1986). *The Morality of Freedom*. Oxford: Clarendon Press.

Raz, J. (1999). 'Agency, Reason and the Good'. In *Engaging Reason: On the Theory of Value and Action*, pp. 22–45. Oxford: Oxford University Press.

Roberts, R. (1988). 'What an Emotion Is: A Sketch'. *The Philosophical Review 97*, 183–209.

Roberts, R. (2003). *Emotions: An Essay in Aid of Moral Psychology*. Cambridge: Cambridge University Press.

Robinson, J. M. (2004). 'Emotion: Biological Fact or Social Construction'. In R. Solomon (Ed.), *Thinking About Feeling: Contemporary Philosophers on Emotion*, pp. 28–42. Oxford: Oxford University Press.

Ryle, G. (1990). *The Concept of Mind*. London: Penguin.

Salmela, M. (2011). 'Can Emotion be Modelled on Perception?'. *Dialectica 65*, 1–29.

Scanlon, T. M. (1999). *What We Owe To Each Other*. Cambridge, MA: Harvard University Press.

Scarantino, A. (2010). 'Insights and Blindspots of the Cognitivist Theory of Emotions'. *British Journal for the Philosophy of Science 61*, 729–68.

Schachter, S. and J. Singer (1962). 'Cognitive, Social, and Physiological Determinants of Emotional State'. *Psychological Review 69*, 379–99.

Schmalz, D. (2012, October 18). 'Climber Alex Honnold on Fear, Fame, and What's Next'. Accessed 28 April 2014, http://adventureblog.nationalgeographic.com/2012/10/18/free-soloist-alex-honnold-on-fear-fame-and-whats-next/.

Schorr, A. (2001). 'Appraisal: The evolution of an idea'. In K. Scherer, A. Schorr, and T. Johnstone (Eds.), *Appraisal Processes in Emotion*, pp. 20–34. Oxford: Oxford University Press.

Schroeder, M. (2008). 'Having Reasons'. *Philosophical Studies 139*, 57–71.

Schwitzgebel, E. (2010). 'Acting Contrary to our Professed Beliefs or The Gulf Between Judgment and Dispositional Belief'. *Pacific Philosophical Quarterly 91*, 531–53.

Schwitzgebel, E. (2014). 'Belief'. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy (Spring 2014 Edition)*. Accessed 21 September 2014, http://plato.stanford.edu/archives/spr2014/entries/belief/.

Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. New York: Cambridge University Press.

Senault, J. (1649). *The Use of the Passions*. London: J. L. Humphrey Moseley.

Shakespeare, W. (1997). *Othello*. The Arden Shakespeare.

Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.

Smith, M. (1998). 'The Possibility of Philosophy of Action'. In J. Bransen and S. E. Cuypers (Eds.), *Human Action, Deliberation, and Causation*, pp. 17–41. Dordrecht: Kluwer Academic Publishers.

Solomon, R. (1976). *The Passions*. Indianapolis: Hackett.

Solomon, R. (2003). 'Against Valence ("Positive" and "Negative" Emotions)'. In R. Solomon (Ed.), *Not Passion's Slave: Emotions and Choice*, pp. 162–77. Oxford: Oxford University Press.

Solomon, R. (2004). 'Emotions, Thoughts, and Feelings'. In R. Solomon (Ed.), *Thinking About Feeling*, pp. 76–88. Oxford: Oxford University Press.

Spinoza, B. (1984). 'Ethics'. In C. Calhoun and R. Solomon (Eds.), *What Is An Emotion? Classic Readings in Philosophical Psychology*. Oxford: Oxford University Press.

Stocker, M. and E. Hegeman (1992). *Valuing Emotions*. Cambridge: Cambridge University Press.

Tappolet, C. (2005). 'Ambivalent Emotions and the Perceptual Account of Emotions'. *Analysis 65*, 229–33.

Tappolet, C. (2010). 'Emotion, Motivation, and Action: The case of fear'. In P. Goldie (Ed.), *The Oxford Handbook of Philosophy of Emotion*, pp. 325–45. Oxford: Oxford University Press.

Tappolet, C. (2012). 'Emotions, Perceptions, and Emotional Illusions'. In C. Calabi (Ed.), *Perceptual Illusions: Philosophical and Psychological Essays*, pp. 205–21. New York: Palgrave Macmillan.

Teroni, F. (2007). 'Emotions and Formal Objects'. *Dialectica 61*, 395–415.

Tolstoy, L. (2001). *War and Peace*. Project Guttenberg. Accessed 17 July 2014, http://www.gutenberg.org/ebooks/2600.

Twain, M. (1885). *The Adventures of Huckleberry Finn*. Charles C. Webster & Co.

Tye, M. (1995). *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.

Wallace, R. J. (1999). 'Three Conceptions of Rational Agency'. *Ethical Theory and Moral Practice 2*, 217–42.

Wallace, R. J. (2014). 'Practical Reason'. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy (Summer 2014 Edition)*. Accessed 1 July 2014, http://plato.stanford.edu/archives/sum2014/entries/practical-reason/.

Way, J. (2009). 'Two Accounts of the Normativity of Rationality'. *Journal of Ethics & Social Philosophy*, 1–8.

Wedgwood, R. (1997). 'The Essence of Response-Dependence'. *European Review of Philosophy 3*, 31–54. Available at http://www-bcf.usc.edu/ wedgwood/respodep.html.

Welsh, I. (1993). *Trainspotting*. Minerva.

Whiting, D. (2012). 'Are Emotions Perceptual Experiences of Value?'. *Ratio 25*, 93–107.

Wiggins, D. (1987). 'A Sensible Subjectivism'. In *Needs, Values, Truth*, pp. 185–211. Oxford: Basil Blackwell.

Wild, M. (2008). 'Critical Reflections on Affective Epistemology'. In G. Brun, U. Doğuoğlu, and D. Kuenzle (Eds.), *Epistemology and Emotions*, pp. 125–48. Aldershot: Ashgate.

Williams, B. (1981). 'Persons, Character, and Morality'. In J. Rachels (Ed.), *Moral Luck: Philosophical Papers 1973-1980*. Cambridge: Cambridge University Press.

Wilson, G. and S. Shpall (2012). 'Action'. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy (Summer 2012 Edition)*. Accessed 10 September 2014, http://plato.stanford.edu/archives/sum2012/entries/action.

Wittgenstein, L. (1958). *Philosophical Investigations*. Oxford: Basil Blackwell.

Wittgenstein, L. and F. Waismann (2003). *Voices of Wittgenstein: The Vienna Circle*. London: Routledge.

Wolf, S. (2012). 'One Thought Too Many: Love, Morality, and the Ordering of Commitment'. In U. Heuer and G. Lang (Eds.), *Luck, Value, and Commitment: Themes form the Ethics of Bernard Williams*. Oxford: Oxford University Press.

Zajonc, R. (1980). 'Feeling and Thinking: Preferences need no inferences'. *American Psychologist 53 (4)*, 151–75.